



Best Local Similarity 65.8%; Pred. No. 4.3e-111;  
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

QY 67 -CAAAPAPF----- 79  
|||  
Db 122 DCVPCPPGHFSPGNQACKPWTNCTLSGKOTRHPASDSLDVACEDRSLLATLLMETORPT 181  
QY 80 -----RSCDKHTHTCCPPAPAEAGAPSVFLFPKPKDT 112  
|||  
Db 182 FRPTTVQSTTVWPRSTSELPSTPTLVEPRSCDKHTHTCCPPAPAEAGAPSVFLFPKPKDT 241  
QY 113 LMSRTEPVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 172  
|||  
Db 242 LMSRTEPVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301  
QY 173 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRREEMTKNOVSLTCLVK 232  
|||  
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVK 361  
QY 233 GFYPDSIAEWESNGQPENNYKTTPPVLDSDGSFGFFLYSKLTVDKSRWQQGNVSCSYMHE 292  
|||  
Db 362 GFYPDSIAEWESNGQPENNYKTTPPVLDSDGSFGFFLYSKLTVDKSRWQQGNVSCSYMHE 421  
QY 293 ALHNHYTQKSLSLSPGK 309  
|||  
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 2  
US-08-494-574-11  
; Sequence 11, Application US/08494574  
; Patent No. 5783665  
; GENERAL INFORMATION:  
; APPLICANT: Baum, Peter  
; APPLICANT: Goodwin, Ray  
; APPLICANT: Fanslow, William  
; APPLICANT: Gayle, Richard  
; TITLE OF INVENTION: No. 5783665el Cytokine Which is a Ligand for  
; TITLE OF INVENTION: OX40  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/494,574  
; FILING DATE: 22-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,827  
; FILING DATE: 23-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia A.  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2806  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-587-0730  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 438 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-494-574-11

Query Match

74.8%; Score 1255; DB 1; Length 438;

Best Local Similarity 65.8%; Pred. No. 4.3e-111;  
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

QY 35 PCSRG-SSWADLDKCMDCAS-----CRARP-----HSDFLG- 66  
|||  
Db 62 PCETGFYNEAVNYDTCKQCTQCNRHSGSELKQNCPTPTQDVTVCRCRPGTQPRQDSGYKLGV 121  
QY 67 -CAAAPAPF----- 79  
|||  
Db 122 DCVPCPPGHFSPGNQACKPWTNCTLSGKOTRHPASDSLDVACEDRSLLATLLMETORPT 181  
QY 80 -----RSCDKHTHTCCPPAPAEAGAPSVFLFPKPKDT 112  
|||  
Db 182 FRPTTVQSTTVWPRSTSELPSTPTLVEPRSCDKHTHTCCPPAPAEAGAPSVFLFPKPKDT 241  
QY 113 LMSRTEPVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 172  
|||  
Db 242 LMSRTEPVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301  
QY 173 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRREEMTKNOVSLTCLVK 232  
|||  
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVK 361  
QY 233 GFYPDSIAEWESNGQPENNYKTTPPVLDSDGSFGFFLYSKLTVDKSRWQQGNVSCSYMHE 292  
|||  
Db 362 GFYPDSIAEWESNGQPENNYKTTPPVLDSDGSFGFFLYSKLTVDKSRWQQGNVSCSYMHE 421  
QY 293 ALHNHYTQKSLSLSPGK 309  
|||  
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 3  
US-09-181-706-8  
; Sequence 8, Application US/09181706  
; Patent No. 6130068  
; GENERAL INFORMATION:  
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,  
; APPLICANT: Robert F. DuBose, Richard S. Johnson  
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
; NUMBER OF INVENTIONS: RECEPTOR DNA AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Janis C. Henry  
; STREET: 51 University St.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/181,706  
; FILING DATE: October 28, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/958,598 (converted to a  
; APPLICATION NUMBER: Provisional, see below)  
; FILING DATE: October 26, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598  
; APPLICATION NUMBER: conversion to Provisional application)  
; FILING DATE: October 26, 1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2631-A  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-181-706-8

Query Match 74.5%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHICPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
DB 34 RSCDKTHICPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 140 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 199  
DB 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 153  
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213  
QY 260 LDSGSEFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTQKSLSLSPGK 309  
DB 214 LDSGSEFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTQKSLSLSPGK 263

## RESULT 4

US-09-458-791-8  
Sequence 8, Application US/09458791  
Patent No. 6174689

GENERAL INFORMATION:  
APPLICANT: Spriggs, Melanie  
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janis C. Henry  
STREET: 51 University St.  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/458,791  
FILING DATE: 10-Dec-1999  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-458-791-8

Query Match 74.5%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHICPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
DB 34 RSCDKTHICPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 140 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 199  
DB 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 153  
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213  
QY 260 LDSGSEFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTQKSLSLSPGK 309  
DB 214 LDSGSEFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTQKSLSLSPGK 263

## RESULT 5

US-09-459-066-8  
Sequence 8, Application US/09459066  
Patent No. 6187909

GENERAL INFORMATION:  
APPLICANT: Spriggs, Melanie  
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janis C. Henry  
STREET: 51 University St.  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/459,066  
FILING DATE:  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-459-066-8

Query Match 74.5%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 RSCDKTHTCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 139  
Db 34 RSCDKTHTCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 93  
Qy 140 YVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGYCKVSNKALPAPIKTI 199  
Db 94 YVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGYCKVSNKALPAPIKTI 153  
Qy 200 KAGQPREQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQPNNTKTPPV 259  
Db 154 KAGQPREQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQPNNTKTPPV 213  
Qy 260 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 309  
Db 214 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 263

## RESULT 6

US-09-180-100-22  
; Sequence 22, Application US/09180100  
; Patent No. 6306395  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, No. 6306395510  
; APPLICANT: NAGATA, Shigekazu  
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
; FILE REFERENCE: 1110-207P  
; CURRENT APPLICATION NUMBER: US/09/180,100  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
; EARLIER FILING DATE: 1997-05-01  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-100-22

Query Match 74.3%; Score 1247; DB 4; Length 376;

Best Local Similarity 65.4%; Pred. No. 2e-110; Matches 249; Conservative 10; Mismatches 32; Indels 90; Gaps 8;

Qy 14 VLGLWAL---LRSA-----GEQ-----APGTAPCSRGS 40  
Db 1 MGIWTLPLVLTSAQNLEGLHHDGFCHPCPPGERKARDCTVNGDEPDCVPCQEGK 60  
Qy 41 SWSADL---DKMDC-----ASCRAP-----HSDFCIGC 67  
Db 61 EYTDKAHFSKRCRCRLCDEGHGLEVEINCTRTQNTKCRCKPNFNCNSTVCEHCDPCTKC 120  
Qy 68 -----AAAPAPRLLWRSCKTHCTPCPPAPAEAGAPSVFLPPK 108  
Db 121 EHGIIKCTLTNTKCKEGRSNEP-----KSCDKTHTCPAPPELLGGPSVFLPPK 175  
Qy 109 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVSVL 168  
Db 176 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVSVL 235  
Qy 169 TVLHODWLNKGYCKVSNKALPAPIKTIKAKGQPREQVYTLPPSREEMTKNOVSLT 228  
Db 236 TVLHODWLNKGYCKVSNKALPAPIKTIKAKGQPREQVYTLPPSREEMTKNOVSLT 295  
Qy 229 CLVKGFPYSDIAVESNGQPNNTKTPPVLDSDGSEFLYSLKLTVDKSRWQGNVFCSC 288  
Db 296 CLVKGFPYSDIAVESNGQPNNTKTPPVLDSDGSEFLYSLKLTVDKSRWQGNVFCSC 355  
Qy 289 VMHEALHNYTKQSLSLSPGK 309  
Db 356 VMHEALHNYTKQSLSLSPGK 376

## RESULT 7

US-09-180-100-11  
; Sequence 11, Application US/09180100  
; Patent No. 6306395  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, No. 6306395510  
; APPLICANT: NAGATA, Shigekazu  
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
; FILE REFERENCE: 1110-207P  
; CURRENT APPLICATION NUMBER: US/09/180,100  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
; EARLIER FILING DATE: 1997-05-01  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-100-11

Query Match 74.1%; Score 1244; DB 4; Length 360;

Best Local Similarity 71.3%; Pred. No. 3.6e-110; Matches 241; Conservative 10; Mismatches 29; Indels 58; Gaps 6;

Qy 24 SVAGEQAPGTAPCSRGSWSADL---DKMDC-----ASCRAP- 59  
Db 29 TVNGDE-PCVPCQEGKEYTDKAHFSKRCRCRLCDEGHGLEVEINCTRTQNTKCRCKPN 87  
Qy 60 -----HSDFCIGC-----AAAPAPRLLWRSCKTHCTPCPC 91  
Db 88 FFCNSTVCEHCDPCTKCEHGIKCTLTNTKCKEGRSNEP-----KSCDKTHTCPC 142  
Qy 92 PAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAK 151  
Db 143 PAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAK 202  
Qy 152 KPREEQYNSTYRVSVTLVHODWLNKGYCKVSNKALPAPIKTIKAKGQPREQVY 211  
Db 203 KPREEQYNSTYRVSVTLVHODWLNKGYCKVSNKALPAPIKTIKAKGQPREQVY 262  
Qy 212 TLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQPNNTKTPPVLDSDGSEFLYSLK 271  
Db 263 TLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQPNNTKTPPVLDSDGSEFLYSLK 322  
Qy 272 LTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 309  
Db 323 LTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 360

## RESULT 8

PCT-US95-03866-12  
; Sequence 12, Application PC/TUS9503866  
; GENERAL INFORMATION:  
; APPLICANT: Cytomed, Inc. (all states except US)  
; APPLICANT: Nocka, Karl (US only)  
; APPLICANT: Lobell, Robert B (US only)  
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND  
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:



Query Match 74.1% Score 1243. DB 5. Length 424.

RESULT 10

```

1  Sequence 3, Application US/08784512
2  Patent No. 5872209
3  GENERAL INFORMATION:
4  APPLICANT: BARTNIK, Eckart
5  APPLICANT: EDENMUeller, Bernd
6  APPLICANT: BUETNER, Frank
7  APPLICANT: CATERSON, Bruce
8  APPLICANT: HUGHES, Clare
9  TITLE OF INVENTION: An artificial
10  TITLE OF INVENTION: and native
11  TITLE OF INVENTION: "AggreCanas"
12  NUMBER OF SEQUENCE: 4
13  CORRESPONDENCE ADDRESS:
14  ADDRESSEE: Foley & Lardner
15  STREET: suite 500, 3000 K Street
16  CITY: Washington, D.C.
17  COUNTRY: USA
18  ZIP: 20007-5109
19  COMPUTER READABLE FORM:
20  MEDIUM TYPE: Floppy disk
21  COMPUTER: IBM PC compatible
22  OPERATING SYSTEM: PC-DOS/MS-DOS
23  SOFTWARE: PatentIn Release #1
24  CURRENT APPLICATION DATA:
25  APPLICATION NUMBER: US/08/784
26  FILING DATE: 17-JAN-1997
27  PRIOR APPLICATION DATA:

```

Sequence 3, Application US/08/84512  
Patent No. 5872209  
GENERAL INFORMATION:  
APPLICANT: BARTNIK, Eckart  
APPLICANT: EIDENMUELLER, Bernd  
APPLICANT: BUETTNER, Frank  
APPLICANT: CATERSON, Bruce  
APPLICANT: HUGHES, Claire  
TITLE OF INVENTION: An artificial recombinant substrate (PAGG 1)  
TITLE OF INVENTION: and native aggrecan to study the proteolytic activity of  
TITLE OF INVENTION: "Aggrecanase" in cell culture systems  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: Suite 500, 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,512  
FILING DATE: 17-JAN-1997  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 96100682.2  
FILING DATE: 18-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: GRANADOS, Patricia D.  
REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 18748/311  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..396  
US-08-784-512-3

Query Match 73.5%; Score 1233.5; DB 2; Length 396;  
Best Local Similarity 83.0%; Pred. No. 4.1e-109;  
Matches 235; Conservative 8; Mismatches 27; Indels 13; Gaps 3;  
QY 27 GEQAGTAPCSRSSWSADLDKCMDCASCARPHSDCLGCAAPAPFLLWRSCDKTH 86  
DB 127 GFPTGLGPATATST-----EDLVVQVTVAVPGQH----LPGGGDPEP-----KSCDKTH 173  
QY 87 TCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 146  
DB 174 TCPPCPAPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 233  
QY 147 HNAKTPREQYNSTYRVWSVLTVLHODWLNKEYCKVSNKALPAPIETISKAKGQPR 206  
DB 234 HNAKTPREQYNSTYRVWSVLTVLHODWLNKEYCKVSNKALPAPIETISKAKGQPR 293  
QY 207 EPQVYILPPSREMTKNQVSLTCLVKGFPSPDI AVEWESNGQPNENYKTTTPVLDSDGSF 266  
DB 294 EPQVYILPPSRDELTKNQVSLTCLVKGFPSPDI AVEWESNGQPNENYKTTTPVLDSDGSF 353  
QY 267 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 309  
DB 354 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 396

RESULT 11  
US-09-176-228-3  
Sequence 3, Application US/09176228  
Patent No. 6180334  
GENERAL INFORMATION:  
APPLICANT: BARTNIK, Eckart  
APPLICANT: EIDENMUELLER, Bernd  
APPLICANT: BUETNER, Frank  
APPLICANT: CATERSON, Bruce  
APPLICANT: HUGHES, Clare  
TITLE OF INVENTION: An artificial recombinant substrate (rAGG 1)  
TITLE OF INVENTION: and native aggregan to study the proteolytic activity of  
TITLE OF INVENTION: "Aggrecanase" in cell culture systems  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: Suite 500, 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/176,228  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,512  
FILING DATE: 17-JAN-1997  
APPLICATION NUMBER: EP 96100682.2  
FILING DATE: 18-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: GRANADOS, Patricia D.  
REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 18748/311  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..396  
US-09-176-228-3  
Query Match 73.5%; Score 1233.5; DB 4; Length 396;  
Best Local Similarity 83.0%; Pred. No. 4.1e-109;  
Matches 235; Conservative 8; Mismatches 27; Indels 13; Gaps 3;  
QY 27 GEQAGTAPCSRSSWSADLDKCMDCASCARPHSDCLGCAAPAPFLLWRSCDKTH 86  
DB 127 GFPTGLGPATATST-----EDLVVQVTVAVPGQH----LPGGGDPEP-----KSCDKTH 173  
QY 87 TCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 146  
DB 174 TCPPCPAPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 233  
QY 147 HNAKTPREQYNSTYRVWSVLTVLHODWLNKEYCKVSNKALPAPIETISKAKGQPR 206  
DB 234 HNAKTPREQYNSTYRVWSVLTVLHODWLNKEYCKVSNKALPAPIETISKAKGQPR 293  
QY 207 EPQVYILPPSREMTKNQVSLTCLVKGFPSPDI AVEWESNGQPNENYKTTTPVLDSDGSF 266  
DB 294 EPQVYILPPSRDELTKNQVSLTCLVKGFPSPDI AVEWESNGQPNENYKTTTPVLDSDGSF 353  
QY 267 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 309  
DB 354 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 396

RESULT 12  
US-08-887-352B-14  
Sequence 14, Application US/08887352B  
Patent No. 5994511  
GENERAL INFORMATION:  
APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
TITLE OF INVENTION: Improving Polypeptides  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,352B  
FILING DATE: 03-Jul-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.  
REGISTRATION NUMBER: 39,044  
REFERENCE/DOCKET NUMBER: P1123  
TELEPHONE: 650/225-1489  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-887-352B-14

Query Match 73.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 1.le-108;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
DB 222 KSCDKTHCPCPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281

QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
DB 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341

QY 200 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPV 259  
DB 342 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPV 401

QY 260 LDSGSEFLYSLKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 309  
DB 402 LDSGSEFLYSLKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 451

RESULT 13  
US-08-887-352B-16  
Sequence 16, Application US/08887352B  
Patent No. 5994511  
GENERAL INFORMATION:  
APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
TITLE OF INVENTION: Improving Polypeptides  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,352B  
FILING DATE: 03-Jul-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.  
REGISTRATION NUMBER: 39,044  
REFERENCE/DOCKET NUMBER: P1123  
TELEPHONE: 650/225-1489

TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-887-352B-16

Query Match 73.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 1.le-108;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
DB 222 KSCDKTHCPCPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281

QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
DB 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341

QY 200 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPV 259  
DB 342 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPV 401

QY 260 LDSGSEFLYSLKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 309  
DB 402 LDSGSEFLYSLKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 451

RESULT 14  
US-08-887-352B-18  
Sequence 18, Application US/08887352B  
Patent No. 5994511  
GENERAL INFORMATION:  
APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
TITLE OF INVENTION: Improving Polypeptides  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,352B  
FILING DATE: 03-Jul-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.  
REGISTRATION NUMBER: 39,044  
REFERENCE/DOCKET NUMBER: P1123  
TELEPHONE: 650/225-1489  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-887-352B-18

Query Match 73.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 1.le-108;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAEAGAPSVLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNW 139  
:|||||  
Db 222 KSCDKTHCPCPAPELLGSPVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNW 281  
:|||||  
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
:|||||  
Db 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341  
:|||||  
QY 200 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259  
:|||||  
Db 342 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 401  
:|||||  
QY 260 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCVVMHEALHNHYTQKSLSLSPGK 309  
:|||||  
Db 402 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCVVMHEALHNHYTQKSLSLSPGK 451  
:|||||

RESULT 15  
US-08-466-151-65  
: Sequence 65, Application US/08466151  
: Patent No. 6037453  
: GENERAL INFORMATION:  
: APPLICANT: Jardieu, Paula M.  
: APPLICANT: Presta, Leonard G.  
: TITLE OF INVENTION: Immunoglobulin Variants  
: NUMBER OF SEQUENCES: 65  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Genentech, Inc.  
: STREET: 1 DNA Way  
: CITY: South San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94080  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Winpatin (Genentech)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/466,151  
: FILING DATE:  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/466163  
: FILING DATE: 06-Jun-1995  
: APPLICATION NUMBER: 08/405617  
: FILING DATE: 15-MAR-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/185899  
: FILING DATE: 26-JAN-1994  
: PRIOR APPLICATION DATA: 07/879495  
: APPLICATION NUMBER: 07-MAY-1992  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 07/744768  
: FILING DATE: 14-AUG-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Svoboda, Craig G.  
: REGISTRATION NUMBER: 39,044  
: REFERENCE/DOCKET NUMBER: P0718P2C1D1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 650/225-1489  
: TELEFAX: 650/952-9881  
: INFORMATION FOR SEQ ID NO: 65:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 451 amino acids  
: TYPE: Amino Acid  
: TOPOLOGY: Linear  
US-08-466-151-65

Best Local Similarity 98.3%; Pred. No. 1.1e-108;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 80 RSCDKTHCPCPAEAGAPSVLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNW 139  
:|||||  
Db 222 KSCDKTHCPCPAPELLGSPVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNW 281  
:|||||  
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
:|||||  
Db 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341  
:|||||  
QY 200 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259  
:|||||  
Db 342 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 401  
:|||||  
QY 260 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCVVMHEALHNHYTQKSLSLSPGK 309  
:|||||  
Db 402 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCVVMHEALHNHYTQKSLSLSPGK 451  
:|||||

Search completed: March 11, 2002, 15:13:12  
Job time: 36 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:13:14 ; Search time 26.47 Seconds  
(without alignments)  
239.740 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_309

Perfect score: 1551

Sequence: 1 EQAPGAPCSRGSWSADLD.....MHEALHHYTKLSLSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PCRTUS.COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description       |
|------------|--------|-------------|--------|----|-------------------|
| 1          | 1255   | 80.9        | 438    | 1  | US-08-097-827-11  |
| 2          | 1255   | 80.9        | 438    | 1  | US-08-494-574-11  |
| 3          | 1250   | 80.6        | 660    | 4  | US-09-181-706-8   |
| 4          | 1250   | 80.6        | 660    | 4  | US-09-458-791-8   |
| 5          | 1250   | 80.6        | 660    | 4  | US-09-459-086-8   |
| 6          | 1243   | 80.1        | 424    | 5  | PCT-US95-03866-12 |
| 7          | 1243   | 80.1        | 424    | 5  | PCT-US95-03866-14 |
| 8          | 1241.5 | 80.0        | 360    | 4  | US-09-180-100-11  |
| 9          | 1241.5 | 80.0        | 376    | 4  | US-09-180-100-22  |
| 10         | 1231.5 | 79.4        | 396    | 2  | US-08-784-512-3   |
| 11         | 1231.5 | 79.4        | 396    | 4  | US-09-176-228-3   |
| 12         | 1230   | 79.3        | 451    | 2  | US-08-887-352B-14 |
| 13         | 1230   | 79.3        | 451    | 2  | US-08-887-352B-16 |
| 14         | 1230   | 79.3        | 451    | 2  | US-08-887-352B-18 |
| 15         | 1230   | 79.3        | 451    | 3  | US-08-466-151-65  |
| 16         | 1230   | 79.3        | 451    | 4  | US-09-109-207C-14 |
| 17         | 1230   | 79.3        | 451    | 4  | US-09-109-207C-16 |
| 18         | 1230   | 79.3        | 451    | 4  | US-09-109-207C-18 |
| 19         | 1230   | 79.3        | 451    | 4  | US-09-282-505-2   |
| 20         | 1230   | 79.3        | 451    | 4  | US-09-054-255-2   |
| 21         | 1230   | 79.3        | 451    | 4  | US-09-296-005-14  |
| 22         | 1230   | 79.3        | 451    | 4  | US-09-296-005-16  |
| 23         | 1230   | 79.3        | 451    | 4  | US-09-296-005-18  |
| 24         | 1230   | 79.3        | 452    | 3  | US-09-027-449-71  |
| 25         | 1230   | 79.3        | 452    | 4  | US-09-026-985-71  |
| 26         | 1230   | 79.3        | 453    | 3  | US-08-466-151-8   |
| 27         | 1230   | 79.3        | 454    | 2  | US-07-934-373C-22 |

|    |        |      |     |   |                   |                   |
|----|--------|------|-----|---|-------------------|-------------------|
| 28 | 1230   | 79.3 | 454 | 3 | US-08-437-642B-22 | Sequence 22, Appl |
| 29 | 1230   | 79.3 | 454 | 5 | PCT-US93-07832-22 | Sequence 22, Appl |
| 30 | 1230   | 79.3 | 467 | 4 | US-09-049-672A-8  | Sequence 8, Appl  |
| 31 | 1230   | 79.3 | 473 | 4 | US-09-049-672A-4  | Sequence 4, Appl  |
| 32 | 1225.5 | 79.0 | 442 | 5 | PCT-US96-10043-9  | Sequence 9, Appl  |
| 33 | 1224   | 78.9 | 232 | 2 | US-08-595-043A-50 | Sequence 50, Appl |
| 34 | 1224   | 78.9 | 331 | 4 | US-09-178-869-2   | Sequence 2, Appl  |
| 35 | 1224   | 78.9 | 371 | 1 | US-08-236-311-7   | Sequence 7, Appl  |
| 36 | 1224   | 78.9 | 371 | 3 | US-08-457-918-7   | Sequence 7, Appl  |
| 37 | 1224   | 78.9 | 437 | 5 | PCT-US96-10043-11 | Sequence 11, Appl |
| 38 | 1224   | 78.9 | 446 | 3 | US-08-397-411-7   | Sequence 7, Appl  |
| 39 | 1224   | 78.9 | 449 | 1 | US-08-458-516-13  | Sequence 13, Appl |
| 40 | 1224   | 78.9 | 459 | 1 | US-08-157-101A-7  | Sequence 7, Appl  |
| 41 | 1224   | 78.9 | 476 | 2 | US-08-378-939-10  | Sequence 10, Appl |
| 42 | 1224   | 78.9 | 476 | 3 | US-08-487-550-4   | Sequence 4, Appl  |
| 43 | 1224   | 78.9 | 476 | 3 | US-08-487-550-12  | Sequence 12, Appl |
| 44 | 1224   | 78.9 | 478 | 3 | US-08-487-550-8   | Sequence 8, Appl  |
| 45 | 1223   | 78.9 | 232 | 3 | US-08-996-139-8   | Sequence 8, Appl  |

#### ALIGNMENTS

RESULT 1  
US-08-097-827-11  
; Sequence 11, Application US/08097827  
; GENERAL INFORMATION:  
; APPLICANT: Baum, Peter  
; Goodwin, Ray  
; Fanslow, William  
; Gayle, Richard  
; TITLE OF INVENTION: Novel Cytokine which is a Ligand for  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 23-Jul-1993  
; APPLICATION NUMBER: US/08/097,827  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia A.  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2806  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-587-0730  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 438 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-08-097-827-11

Query Match 80.9%; Score 1255; DB 1; Length 438;  
Best Local Similarity 65.8%; Pred. No. 7.8e-113;  
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;  
QY 8 PCSRG-SWSADLDKCMDCAS-----CRARP-----HSDFCLG- 39  
||| : : : : : ||| : : : : :  
DB 62 PCETGFYNAVNYDKQCTQCNRHSRSGSELKQNCPTQDTVCRCRGTGTPRODSGYKLG 121

QY 40 -CAAAPAPF-----RLLW----- 52  
Db 122 DCVPCPGHSPGNQACKPWTNCTLSGKQTRHPASDSDAVCEDRSLATLLWETORPT 181  
QY 53 -----RSCDKHTHTCCPCPAPEAGAPSVFLFPPKPKDT 85  
Db 182 FRPTTVQSTTVWPTSELPSTPTLVPRSCDKHTHTCCPCPAPEAGAPSVFLFPPKPKDT 241  
QY 86 LMSRTPETVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 145  
Db 242 LMSRTPETVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301  
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNOVSLTCLVK 205  
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 361  
QY 206 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSYMHE 265  
Db 362 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSYMHE 421  
QY 266 ALHNHYTQKSLSLSPGK 282  
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 2  
US-08-494-574-11  
; Sequence 11, Application US/08494574  
; Patent No. 5783665  
; GENERAL INFORMATION:  
; APPLICANT: Baum, Peter  
; APPLICANT: Goodwin, Ray  
; APPLICANT: Fauslow, William  
; APPLICANT: Gayle, Richard  
; TITLE OF INVENTION: No. 5783665el Cytokine Which is a Ligand for  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/494,574  
; FILING DATE: 22-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,827  
; FILING DATE: 23-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia A.  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2806  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-587-0730  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 438 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-494-574-11

Query Match

80.9%; Score 1255; DB 1; Length 438;

Best Local Similarity 65.8%; Pred. No. 7.8e-113;  
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;  
QY 8 PCSRG-SSWADLDKCMDCAS-----CRARP-----HSDFLG- 39  
Db 62 PCETGFYNEAVNYDTCKQCTQCNRHSGSELKQNCPTPTQDTCRCRPGTOPRODSGYKLGV 121  
QY 40 -CAAAPAPF-----RLLW----- 52  
Db 122 DCVPCPGHSPGNQACKPWTNCTLSGKQTRHPASDSDAVCEDRSLATLLWETORPT 181  
QY 53 -----RSCDKHTHTCCPCPAPEAGAPSVFLFPPKPKDT 85  
Db 182 FRPTTVQSTTVWPTSELPSTPTLVPRSCDKHTHTCCPCPAPEAGAPSVFLFPPKPKDT 241  
QY 86 LMSRTPETVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 145  
Db 242 LMSRTPETVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301  
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNOVSLTCLVK 205  
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 361  
QY 206 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSYMHE 265  
Db 362 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSYMHE 421  
QY 266 ALHNHYTQKSLSLSPGK 282  
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 3  
US-09-181-706-8  
; Sequence 8, Application US/09181706  
; Patent No. 6130068  
; GENERAL INFORMATION:  
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,  
; APPLICANT: Robert F. DuBose, Richard S. Johnson  
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janis C. Henry  
; STREET: 51 University St.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/181,706  
; FILING DATE: October 28, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/958,598 (converted to a  
; APPLICATION NUMBER: Provisional, see below)  
; FILING DATE: October 28, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598  
; APPLICATION NUMBER: conversion to Provisional application)  
; FILING DATE: October 26, 1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2631-A  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-181-706-8

Query Match 80.6%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 4.2e-112;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112  
DB 34 RSCDKTHTCPPCPAPEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 113 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172  
DB 94 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153  
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 213  
QY 233 LSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKLSLSLSPGK 282  
DB 214 LSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKLSLSLSPGK 263

## RESULT 4

US-09-458-791-8  
Sequence 8, Application US/09458791  
Patent No. 6174689  
GENERAL INFORMATION:  
APPLICANT: Spriggs, Melanie  
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
RECEPTOR DNA AND POLYPEPTIDES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janis C. Henry  
STREET: 51 University St.  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
FILING DATE: 10-Dec-1999  
APPLICATION NUMBER: US/09/458,791  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-458-791-8

Query Match 80.6%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 4.2e-112;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112  
DB 34 RSCDKTHTCPPCPAPEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 113 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172  
DB 94 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153  
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 213  
QY 233 LSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKLSLSLSPGK 282  
DB 214 LSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKLSLSLSPGK 263

## RESULT 5

US-09-459-066-8  
Sequence 8, Application US/09459066  
Patent No. 6187909  
GENERAL INFORMATION:  
APPLICANT: Spriggs, Melanie  
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
RECEPTOR DNA AND POLYPEPTIDES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janis C. Henry  
STREET: 51 University St.  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/459,066  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-459-066-8

Query Match 80.6%; Score 1250; DB 4; Length 660;  
Best Local Similarity 100.0%; Pred. No. 4.2e-112;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





Db 265 KTKPREEQNSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIEKTISKAKGQPREPQ 324  
QY 183 VYTLPPSREMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 242  
Db 325 VYTLPPSREMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 384  
QY 243 SKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282  
Db 385 SKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 424

## RESULT 8

US-09-180-100-11  
; Sequence 11, Application US/09180100  
; Patent No. 6306395  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, No. 630639510  
; APPLICANT: NAGATA, Shigekazu  
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
; FILE REFERENCE: 1110-207P  
; CURRENT APPLICATION NUMBER: US/09/180,100  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
; EARLIER FILING DATE: 1997-05-01  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-100-11

Query Match 80.0%; Score 1241.5; DB 4; Length 360;  
Best Local Similarity 72.2%; Pred. No. 1.2e-111;  
Matches 239; Conservative 7; Mismatches 28; Indels 57; Gaps 5;

QY 4 PGTAPCSRGSSWSADL---DKCMDC-----ASCRARP----- 32  
Db 35 PDCVPCQEGKEYTDKAHFSSKRCRCLDEGHGLEVEINCTRTQNTKCRCKPNFFCNSTV 94  
QY 33 --HSDPCLGC-----AAAPPAPFRLLRSCDKTHTCPCPAPEAG 71  
Db 95 CEHCDPCTKCEHGIKECTLTNTKCKEGRSNEP-----KSCDKTHTCPCPAPELIG 149  
QY 72 APSVFLFPPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPRREQY 131  
Db 150 GPSVFLFPPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPRREQY 209  
QY 132 NSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 191  
Db 210 NSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 269  
QY 192 EMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 251  
Db 270 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 329  
QY 252 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282  
Db 330 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 360

## RESULT 9

US-09-180-100-22  
; Sequence 22, Application US/09180100  
; Patent No. 6306395  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, No. 630639510  
; APPLICANT: NAGATA, Shigekazu  
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
; FILE REFERENCE: 1110-207P  
; CURRENT APPLICATION NUMBER: US/09/180,100  
; CURRENT FILING DATE: 1998-11-02

; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
; EARLIER FILING DATE: 1997-05-01  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-100-22

Query Match 80.0%; Score 1241.5; DB 4; Length 376;  
Best Local Similarity 72.2%; Pred. No. 1.2e-111;  
Matches 239; Conservative 7; Mismatches 28; Indels 57; Gaps 5;

QY 4 PGTAPCSRGSSWSADL---DKCMDC-----ASCRARP----- 32  
Db 51 PDCVPCQEGKEYTDKAHFSSKRCRCLDEGHGLEVEINCTRTQNTKCRCKPNFFCNSTV 110  
QY 33 --HSDPCLGC-----AAAPPAPFRLLRSCDKTHTCPCPAPEAG 71  
Db 111 CEHCDPCTKCEHGIKECTLTNTKCKEGRSNEP-----KSCDKTHTCPCPAPELIG 165  
QY 72 APSVFLFPPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPRREQY 131  
Db 166 GPSVFLFPPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPRREQY 225  
QY 132 NSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 191  
Db 226 NSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 285  
QY 192 EMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 251  
Db 286 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 345  
QY 252 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282  
Db 346 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 376

## RESULT 10

US-08-784-512-3  
; Sequence 3, Application US/08784512  
; Patent No. 5872209  
; GENERAL INFORMATION:  
; APPLICANT: BARTNIK, Eckart  
; APPLICANT: EIDENMUELLER, Bernd  
; APPLICANT: BUETTNER, Frank  
; APPLICANT: CATERSON, Bruce  
; APPLICANT: HUGHES, Clare  
; TITLE OF INVENTION: An artificial recombinant substrate (rAGG 1)  
; TITLE OF INVENTION: and native aggregan to study the proteolytic activity of  
; TITLE OF INVENTION: "Aggrecanase" in cell culture systems  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: Suite 500, 3000 K Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,512  
; FILING DATE: 17-JAN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 96100682.2  
; FILING DATE: 18-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 18748/311  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:

NAME/KEY: Protein  
LOCATION: 1..396  
US-08-784-512-3

Query Match 79.4%; Score 1231.5; DB 2; Length 396;

Best Local Similarity 83.9%; Pred. No. 1.2e-110; Mismatches 24; Indels 13; Gaps 3;

Matches 234; Conservative 8; Mismatches 24; Indels 13; Gaps 3;

QY 4 PGAPCRSGSSWADLDKCMDCASCRAPHSDFCLGCAAAAPPAPFRLLRSCDKTHTCPP 63  
DB 131 PGLGPATAFTS----EDLVQVTA VPGPH---LPGGDPPEP-----KSCDKTHTCPP 177  
QY 64 CPAPEAGASVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 123  
DB 178 CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 237  
QY 124 TKPREEQNSTYRVVSVTLVHLDWLNKGEYCKKVKSNKALPAPLEKTSKAKGPREPQV 183  
DB 238 TKPREEQNSTYRVVSVTLVHLDWLNKGEYCKKVKSNKALPAPLEKTSKAKGPREPQV 297  
QY 184 YTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFELY 243  
DB 298 YTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFELY 357  
QY 244 KLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 282  
DB 358 KLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 396

#### RESULT 11

US-09-176-228-3  
Sequence 3, Application US/09176228  
Patent No. 6180334

GENERAL INFORMATION:

APPLICANT: BARTNIK, Eckart  
APPLICANT: EIDENMUELLER, Bernd  
APPLICANT: BUETTNER, Frank  
APPLICANT: CATERSON, Bruce  
APPLICANT: HUGHES, Clare

TITLE OF INVENTION: An artificial recombinant substrate (RAGG 1)  
TITLE OF INVENTION: and native aggregan to study the proteolytic activity of  
TITLE OF INVENTION: "Aggrecanase" in cell culture systems

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner  
STREET: Suite 500, 3000 K Street, N.W.  
CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/176,228

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/784,512  
FILING DATE: 17-JAN-1997  
APPLICATION NUMBER: EP 96100682.2  
FILING DATE: 18-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 18748/311

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 396 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..396

US-09-176-228-3

Query Match 79.4%; Score 1231.5; DB 4; Length 396;

Best Local Similarity 83.9%; Pred. No. 1.2e-110; Mismatches 24; Indels 13; Gaps 3;

Matches 234; Conservative 8; Mismatches 24; Indels 13; Gaps 3;

QY 4 PGAPCRSGSSWADLDKCMDCASCRAPHSDFCLGCAAAAPPAPFRLLRSCDKTHTCPP 63

DB 131 PGLGPATAFTS----EDLVQVTA VPGPH---LPGGDPPEP-----KSCDKTHTCPP 177

QY 64 CPAPEAGASVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 123

DB 178 CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 237

QY 124 TKPREEQNSTYRVVSVTLVHLDWLNKGEYCKKVKSNKALPAPLEKTSKAKGPREPQV 183

DB 238 TKPREEQNSTYRVVSVTLVHLDWLNKGEYCKKVKSNKALPAPLEKTSKAKGPREPQV 297

QY 184 YTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFELY 243

DB 298 YTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFELY 357

QY 244 KLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 282

DB 358 KLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 396

#### RESULT 12

US-08-887-352B-14

Sequence 14, Application US/08887352B

Patent No. 5994511

GENERAL INFORMATION:

APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe

TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of

TITLE OF INVENTION: Improving Polypeptides

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/887,352B

;  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-14

Query Match 79.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 2.1e-110;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 53 RSCDKTHTCPPCPAPEAGAPSVFLPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 112  
Db 222 KSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 281  
  
QY 113 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVQLHQLWLNKCKYKCKVSNKALPAPIEKTIS 172  
Db 282 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVQLHQLWLNKCKYKCKVSNKALPAPIEKTIS 341  
  
QY 173 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 232  
Db 342 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 401  
  
QY 233 LSDSGSFYLYSLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 282  
Db 402 LSDSGSFYLYSLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 451

RESULT 13  
US-08-887-352B-16  
; Sequence 16, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe  
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
; OPERATING SYSTEM: Improving Polypeptides  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids

;  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-16

Query Match 79.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 2.1e-110;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 53 RSCDKTHTCPPCPAPEAGAPSVFLPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 112  
Db 222 KSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 281  
  
QY 113 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVQLHQLWLNKCKYKCKVSNKALPAPIEKTIS 172  
Db 282 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVQLHQLWLNKCKYKCKVSNKALPAPIEKTIS 341  
  
QY 173 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 232  
Db 342 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 401  
  
QY 233 LSDSGSFYLYSLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 282  
Db 402 LSDSGSFYLYSLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 451

RESULT 14  
US-08-887-352B-18  
; Sequence 18, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe  
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
; OPERATING SYSTEM: Improving Polypeptides  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-18

Query Match 79.3%; Score 1230; DB 2; Length 451;  
Best Local Similarity 98.3%; Pred. No. 2.1e-110;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 53 RSCDKTHTCPPCPAPEAGAPSVFLPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 112  
Db 222 KSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 281

Qy 113 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 172  
Db 282 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 341  
Qy 173 KAKQPREPOQVYITLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTTPPV 232  
Db 342 KAKQPREPOQVYITLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTTPPV 401  
Qy 233 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 282  
Db 402 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 451

RESULT 15  
US-08-466-151-65  
; Sequence 65, Application US/08466151  
; Patent No. 6037453  
; GENERAL INFORMATION:  
; APPLICANT: Jardiou, Paula M.  
; APPLICANT: Presta, Leonard G.  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,151  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/466163  
; FILING DATE: 06-Jun-1995  
; APPLICATION NUMBER: 08/405617  
; FILING DATE: 15-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/185899  
; FILING DATE: 26-JAN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/879495  
; FILING DATE: 07-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/744768  
; FILING DATE: 14-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P0718P2C1D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; US-08-466-151-65

Query Match 79.38; Score 1230; DB 3; Length 451;  
Best Local Similarity 98.38; Pred. No. 2.le-110;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 53 RSCDKTHTCPCPAPEAGAPSVLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

Db 222 KSCDKTHTCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281  
Qy 113 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 172  
Db 282 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 341  
Qy 173 KAKQPREPOQVYITLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTTPPV 232  
Db 342 KAKQPREPOQVYITLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTTPPV 401  
Qy 233 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 282  
Db 402 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 451

Search completed: March 11, 2002, 15:13:15  
Job time: 39 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:14 ; Search time 53.92 Seconds

(without alignments)  
387,401 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_309

Perfect score: 1551

Sequence: 1 EQAPGAPCSRGSSWEADLD.....MHEALHNYTKSLSPGK 282

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Query Score | Match | Length | ID          | Description         |
|------------|-------------|-------|--------|-------------|---------------------|
| 1          | 1551        | 100.0 | 309    | 22 AAU03500 | Human TWEAKR recept |
| 2          | 1258        | 81.1  | 451    | 20 AAY43139 | NAIL-Fc protein se  |
| 3          | 1255        | 80.9  | 438    | 16 AAR81882 | Plasmid pDC406/OX4  |
| 4          | 1255        | 80.9  | 438    | 19 AAW48976 | OX40/Fc mutein. C   |
| 5          | 1250        | 80.6  | 660    | 21 AAY13463 | Ectromelia A39R se  |
| 6          | 1250        | 80.6  | 660    | 20 AAY13463 | Ectromelia A39R se  |
| 7          | 1250        | 80.6  | 660    | 21 AAB28523 | Ectromelia A39R se  |
| 8          | 1250        | 80.6  | 660    | 22 AAB70132 | Ectromelia A39R se  |
| 9          | 1247        | 80.4  | 450    | 20 AAB31694 | Amino acid sequenc  |
| 10         | 1247        | 80.4  | 453    | 20 AAY15239 | ULBP1-Fc fusion po  |
| 11         | 1243        | 80.1  | 424    | 16 AAW14764 | ULBP2-Fc fusion po  |
|            |             |       |        |             | Human soluble kit   |

|    |        |      |     |    |           |                    |
|----|--------|------|-----|----|-----------|--------------------|
| 12 | 1243   | 80.1 | 424 | 16 | AAW14765  | Human soluble kit  |
| 13 | 1241.5 | 80.0 | 376 | 19 | AAW60037  | Antigenic peptide  |
| 14 | 1239.5 | 79.9 | 375 | 18 | AAW06683  | Heregulin-alpha fu |
| 15 | 1239   | 79.9 | 535 | 20 | AAY17414  | SYPH1-26 disintegr |
| 16 | 1235.5 | 79.7 | 234 | 18 | AAW16686  | Human IgG1 FC port |
| 17 | 1235   | 79.6 | 400 | 21 | AAY15123  | Porcine CTLA-4-Ig  |
| 18 | 1231.5 | 79.4 | 396 | 18 | AAW18574  | Aggrecanase artifi |
| 19 | 1231.5 | 79.4 | 396 | 18 | AAW18575  | Aggrecanase artifi |
| 20 | 1230.5 | 79.3 | 488 | 21 | AAI97175  | Human FGF-R1 Extra |
| 21 | 1230   | 79.3 | 232 | 22 | AAE02642  | Human immunoglobul |
| 22 | 1230   | 79.3 | 292 | 22 | AAE02646  | Human immunoglobul |
| 23 | 1230   | 79.3 | 292 | 22 | AAE02647  | Human IgG1 FC regi |
| 24 | 1230   | 79.3 | 330 | 20 | AAE050153 | Human heavy chain  |
| 25 | 1230   | 79.3 | 339 | 22 | AAE02648  | Human immunoglobul |
| 26 | 1230   | 79.3 | 401 | 18 | AAW10537  | Leptin 1-167/IgG1  |
| 27 | 1230   | 79.3 | 444 | 21 | AAI32263  | Humanised anti-CD2 |
| 28 | 1230   | 79.3 | 451 | 20 | AAI50031  | Human E27 anti-IgE |
| 29 | 1230   | 79.3 | 451 | 20 | AAW95659  | Mus musculus anti- |
| 30 | 1230   | 79.3 | 451 | 20 | AAW95661  | Mus musculus anti- |
| 31 | 1230   | 79.3 | 451 | 20 | AAW95663  | Mus musculus anti- |
| 32 | 1230   | 79.3 | 451 | 21 | AAW07473  | Amino acid sequenc |
| 33 | 1230   | 79.3 | 451 | 21 | AAI5201   | Light chain amino  |
| 34 | 1230   | 79.3 | 451 | 22 | AAW74212  | E27 anti-IgE antib |
| 35 | 1230   | 79.3 | 451 | 22 | AAW47088  | Anti-IgE antibody, |
| 36 | 1230   | 79.3 | 451 | 22 | AAW76948  | Full length heavy  |
| 37 | 1230   | 79.3 | 451 | 22 | AAW78950  | Full length heavy  |
| 38 | 1230   | 79.3 | 451 | 22 | AAW78952  | Full length heavy  |
| 39 | 1230   | 79.3 | 452 | 19 | AAW69316  | Anti-IL-8 humanise |
| 40 | 1230   | 79.3 | 452 | 20 | AAI29458  | Recombinant immuno |
| 41 | 1230   | 79.3 | 452 | 21 | AAI30322  | Humanised anti-IL- |
| 42 | 1230   | 79.3 | 452 | 21 | AAI77666  | Humanised anti-IL- |
| 43 | 1230   | 79.3 | 453 | 14 | AAI33311  | Humanised MAb11 Ve |
| 44 | 1230   | 79.3 | 453 | 20 | AAI50151  | Antibody F19 chime |
| 45 | 1230   | 79.3 | 453 | 21 | AAI85199  | Heavy chain amino  |

#### ALIGNMENTS

RESULT 1

AAU03500

ID AAU03500 standard; Protein; 309 AA.

AC AAU03500;

XX 26-SEP-2001 (first entry)

DT Human TWEAKR receptor-Fc (TWEAKR-Fc) fusion polypeptide.

XX TWEAKR receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;  
XX ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
XX retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
XX rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
XX corneal graft neovascularisation; psoriasis; metastatic condition;  
XX malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
XX preneoplastic condition; myocardial angiogenesis; wound granulation;  
XX scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
XX atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
XX peripheral atherosclerosis; human IgG1; TWEAKR-Fc; fusion protein.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..27

FT Protein /note= "Signal peptide"

FT Protein 28..309

FT Domain /note= "Mature human TWEAKR-Fc fusion protein."

FT Region /note= "Specifically referred to in Claim 11"

FT Region /note= "From TWEAKR extracellular domain. Specifically referred to in Claims 4 and 10"

FT Region 80..81

| FT | Region  | /note= "From a BgIII cloning site" |
|----|---|------------------------------------|
| FT | 82..309   |                                    |
| FT |   | /note= "Fc portion"                |
| XX | WC0200145730-A2..   |                                    |
| XX | 28-JUN-2001..   |                                    |
| XX | 19-DEC-2000; 2000WO-US34755..   |                                    |
| XX | 20-DEC-1999; 99US-0172878..   |                                    |
| XX | 10-MAY-2000; 2000US-0203347..   |                                    |
| XX | (IMMV ) IMMUNEX CORP..  |                                    |
| XX | Wiley SR;   |                                    |
| XX | WPI; 2001-417975/44..   |                                    |
| XX | N-PSDB; AAS03965..  |                                    |
| PT | Modulating angiogenesis in a mammal for treating diseases mediated by     |                                    |
| PT | angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  |                                    |
| PT | peripheral tissue, by administering antagonist or agonist of TWEAK        |                                    |
| XX | receptor  |                                    |
| XX | Claim 4; Page 45-46; 46pp; English.                                       |                                    |
| XX | The sequence represents a fusion protein consisting of the human TWEAK    |                                    |
| XX | receptor (TWEAKR) protein extracellular domain fused to an Fc portion     |                                    |
| XX | from human IgG1. This fusion protein, TWEAKR-Fc, is used in the           |                                    |
| XX | preparation of TWEAKR agonists and antagonists. The TWEAK protein is a    |                                    |
| XX | member of the tumour necrosis factor (TNF) family and induces             |                                    |
| XX | angiogenesis. TWEAKR may therefore be used to screen for and develop      |                                    |
| XX | TWEAKR agonists and antagonists for the modulation of angiogenesis, to be |                                    |
| XX | used in the treatment and diagnosis of human disease. The disorders       |                                    |
| XX | mediated by angiogenesis include ocular disorders characterised by ocular |                                    |
| XX | neovascularisation such as diabetic retinopathy, neovascular glaucoma,    |                                    |
| XX | retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,      |                                    |
| XX | rubeosis, uveitis, macular degeneration and corneal graft                 |                                    |
| XX | neovascularisation, and inflammatory diseases such as arthritis,          |                                    |
| XX | rheumatism and psoriasis. Other treatable diseases include malignant and  |                                    |
| XX | metastatic conditions such as sarcomas and carcinomas, benign tumours and |                                    |
| XX | preneoplastic conditions, myocardial angiogenesis, haemophilic joints,    |                                    |
| XX | sceleroderma, vascular adhesions, atherosclerotic plaque                  |                                    |
| XX | neovascularisation, telangiectasia, wound granulation, coronary           |                                    |
| XX | atherosclerosis, peripheral atherosclerosis and ischaemia.                |                                    |
| XX | Sequence 309 AA;  |                                    |
| QY | Query Match 100.08; Score 1551; DB 22; Length 309;                        |                                    |
| QY | Best Local Similarity 100.08; Pred. No. 1.7e-111;                         |                                    |
| QY | Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0               |                                    |
| QY | 1 EQAPGTAPCSRGSSNSADLDKDCMCASCARAPSHSDFCLGCAAAAPPAPFRLLRSCDKTHT 60        |                                    |
| Db | 28 eqapgtapcsrgssnsadldkdcmcascrarphsdfclgcaaaappapfllrwscktht 87         |                                    |
| QY | 61 CPPCPAPEAGAPSVFLFPFKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVEH 120         |                                    |
| Db | 88 cppcpapeagapsvflfppkpkdtlmisrtpcvtdvshedpevkfnwydgvveh 147             |                                    |
| QY | 121 NAKTKPREQYNSTRVYSVLTVLHQDWLNGKEYCKVSKNALPAPTEKTIISKAKGP 180           |                                    |
| Db | 148 naktkpreeqynstrvysvltvlhqdwlngkeyckvsknalpaplektiskakgpre 207         |                                    |
| QY | 181 PQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWEWSNGOPENNYKTTTPVLDSDG 240         |                                    |
| Db | 208 pqvyltppreemtknqvsltlcvkgfypsdiavewesngopennykttppvldsdgsff 267       |                                    |
| QY | 241 LYSKLTVDKSRWQOQNVFSCSYMHEALHNHYTQKSLSLSPGK 282                        |                                    |
| Db | 268 lvsyltdvksrwgqanvfscsymhealnbhvtqkslsispak 309                        |                                    |

## RESULT 2

AAY43139  
 ID AAY43139 standard; Protein; 451 AA.  
 XX AC AAY43139;  
 XX DT 23-DEC-1999 (first entry)  
 XX DE NAIL-Fc protein sequence.  
 XX KW NAIL; human; natural killer cell; NK Cell Activation Inducing Ligand;  
 KW immunogen; B cell stimulation; cancer cell proliferation; inhibitor;  
 KW CD48; dendritic cell; cytotoxic T cell; cancer; autoimmune disease;  
 KW therapy; fusion protein.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 OS XX  
 PN WO9950297-A1.  
 XX PD 07-OCT-1999.  
 XX 23-MAR-1999; 99WO-US062115.  
 XX PF 27-MAR-1998; 98US-0079845.  
 XX PR 17-AUG-1998; 98US-0096750.  
 XX PA (IMMV ) IMMUNEX CORP.  
 XX PI Kubin MZ, Goodwin RG;  
 XX WFI; 1999-591275/50.  
 XX  
 PT New NK cell activation ligand (NAIL) polypeptides stimulate B, NK and T  
 PT cells and eliminate cancer cells -  
 PS Disclosure; Page 25-26; 113pp; English.  
 PS  
 XX This sequence is a fusion protein containing the human NK (Natural  
 CC Killer) Cell Activation Inducing Ligand (NAIL) protein of the invention.  
 CC An immunogenic composition containing a NAIL polypeptide may be  
 CC useful for: stimulating B cells; stimulating NK cells; inhibiting the  
 CC proliferation of cancer cells; chelating soluble CD48 in a patient;  
 CC inhibiting the binding of CD48 with NAIL on the cell surface in a  
 CC patient; inhibiting the binding of NAIL with CD48 on the cell surface in  
 CC a patient; stimulating dendritic cells in a patient; stimulating NK cells  
 CC in a patient; stimulating cytotoxic T cells in a patient; inhibiting the  
 CC stimulation of NK cells or cytotoxic T cells in a patient; inhibiting the  
 CC stimulation of B cells or dendritic cells in a patient. The nucleic acids  
 CC of the invention can be used as probes to identify nucleic acid encoding  
 CC proteins having NAIL activity and to inhibit expression of polypeptides  
 CC encoded by the NAIL gene. Another use of the polypeptide is as a research  
 CC tool for studying the biological effects that result from inhibiting  
 CC NAIL/CD48 interactions on different cell types and to study cell signal  
 CC transduction. Antibodies against NAIL can be used to detect the presence  
 CC of NAIL in a sample and can be used therapeutically to bind to NAIL and  
 CC inhibit its activity in vivo. The antibodies can also be useful in the  
 CC diagnosis of pathological states that result in overexpression or  
 CC underexpression of NAIL such as in cancers and autoimmune diseases. The  
 CC invention provides polypeptides and antibodies suitable for use in  
 CC studies of the modulation of NK, T cell and B cell activity and in the  
 CC selection of specific cell types. Furthermore, it provides polypeptides  
 CC and antibodies suitable for use in studies of CD48, where a lack of  
 CC suitable reagents have previously been unavailable.  
 XX Sequence 451 AA;

Query Match 81.1%; Score 1258; DB 20; Length 451;  
Best Local Similarity 86.3%; Pred. No. 8.8e-89;  
Matches 240: Conservative 3; Mismatches 15; Indels 20

```
QY 6 TAPCRGSSW-SADLDKCMDCASCRARPHSDFCLGCAAPAPFRLMPSCDKTHTCPPC 64
Db 193 tcnvsnpsweshntlnltqdcna---hgef-----rrscdkthtcppc 233
QY 65 PAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT 124
Db 234 papeagapsvflfpkpkdtlmisrtpevtcvvdshedpevkfnwydvevhnakt 293
QY 125 KPREGQYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVY 184
Db 294 kpreeqynstyrvsvltvltlhdwlngkeykckvsnkalpapiektiskakgprepgvy 353
QY 185 TLPPSRREMTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLXSK 244
Db 354 tlppsrreemtknovsltclvkgyfypsdiavewesngqpennnykttppvldsdgsfflxsk 413
QY 245 LVVDKSRMQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 282
Db 414 ltvdkrsrqggnvfscsvmhealhnhytqkslsispkg 451

RESULT 3
AAR81882
ID AAR81882 standard; Protein; 438 AA.
XX AC AAR81882;
XX DT 30-MAR-1996 (first entry)
XX DE Plasmid pBC406/OX40/Fc* encoding an OX40/Fc mutein protein.
XX KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
XX KW pDC406/OX40/Fc*; membrane glycoprotein.
XX OS Synthetic.
XX PN US5457035-A.
XX PD 10-OCT-1995.
XX PF 23-JUL-1993; 93US-0097827.
XX PR 23-JUL-1993; 93US-0097827.
XX PA (IMMV ) IMMUNEX CORP.
XX PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX WPI: 1995-357992/46.
XX DR N-PSDB; AAT00829.
XX PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
XX PT and host cells, used to produce recombinant ligand used in e.g.
XX PT prim. T cell culture, to modulate immune response etc.
XX PS Example 2; Column 35-38; 26pp; English.
XX CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
XX CC and is used to express a soluble OX40/Fc mutein fusion protein for
XX CC use in detecting cDNA clones encoding a OX40 ligand. The Fc
XX CC fragment may be derived from human IgG1, and the plasmid may be
XX CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
XX CC line. Culture supernatant was purified by affinity chromatography
XX CC and this was used, together with labeled goat anti-human IgG to
XX CC screen various cell lines.
XX SQ Sequence 438 AA;
```

Query Match 80.9%; Score 1255; DB 16; Length 438;  
Best Local Similarity 65.8%; Pred. No. 1.4e-88;  
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

```
QY 8 PCSRG-SSWSADLDKCMDCAS-----CRARP-----HSDFCLG- 39
Db 62 pceetfyneavnydtckgctqcnhrsgselkqncptqdtvcrpqtgrqdgsgyklgv 121
QY 40 -CAAAPAPAPF-----RSCDKTHTCPPCPAPEAGAPSVFLFPKPKDT 85
Db 122 dcvcppghfsgnngackpwtncnctlsqkqtrhpasdsldavcedrslatlilwetqrpt 181
QY 53 -----RSCDKTHTCPPCPAPEAGAPSVFLFPKPKDT 85
Db 182 frptvtqvstvtwprtselpstptllveprscdkthtcppcpapeagapsvflfpkpkdt 241
QY 86 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVSVLTVLH 145
Db 242 lmisrtpevtcvvdvshedpevkfnwydvevhnaktkpreeqynstyrvsvltvln 301
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRREMTKNOVSLTCLVK 205
Db 302 qdwlngkeykckvsnkalpapiektiskakgprepgvytlppsrdelcknqvsclclvk 361
QY 206 GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQGNVFCSCVMHE 265
Db 362 gfypsdiavewesngqpennnykttppvldsdgsfflyskltvdksrwgqgnvfscsvmhe 421
QY 266 ALHNHYTQKSLSLSPGK 282
Db 422 alhnhytqkslsispkg 438

RESULT 4
AAW48976
ID AAW48976 standard; Protein; 438 AA.
XX AC AAW48976;
XX DT 25-SEP-1998 (first entry)
XX DE OX40/Fc mutein.
XX KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;
XX KW chimeric.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Mus sp.
XX FH Key Location/Qualifiers
XX FT Region 1..206
XX FT /note= "Extracellular domain of mouse OX40"
XX FT Region 207..438
XX FT /note= "Mutant Fc region of human IgG1 antibody"
XX FT Misc-difference 225
XX FT /note= "changed from Leu in wild-type to Ala in
XX FT mutant"
XX FT Misc-difference 226
XX FT /note= "changed from Leu in wild-type to Gly in
XX FT mutant"
XX FT Misc-difference 228
XX FT /note= "changed from Gly in wild-type to Ala in
XX FT mutant"
XX PN US5783665-A.
XX PD 21-JUL-1998.
XX PF 22-JUN-1995; 95US-0494574.
XX PR 23-JUL-1993; 93US-0097827.
XX PR 22-JUN-1995; 95US-0494574.
XX PA (IMMV ) IMMUNEX CORP.
XX PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
```

```
XX WPI; 1998-427099/36.
DR N-PSDB; AAV32636.
XX
XX Purified polypeptide OX-40 ligands - for co-stimulation of T-cell
PT production and binding assays for OX-40 and homologues
XX
XX Example 2; Col 37-40; 26pp; English.
XX
XX The present sequence represents the OX40/Fc fusion protein that
CC contains the extracellular domain of mouse OX40 fused to the mutated
CC FC region of the human IgG1 antibody. The fusion protein was used
CC for detecting cDNA clones encoding an OX40 ligand. The invention
CC claims for a murine OX40-L cytokine (AAW48975) that binds to the murine
CC T cell antigen, OX40. The OX40-L protein is claimed to be useful for
CC co-stimulation of T-cell production and in binding assays for
CC detecting OX40 or its homologues. The OX40-L protein is also claimed
CC to generate a TH-2 immune response.
XX
XX Sequence 438 AA;
XX
XX Query Match 80.9%; Score 1255; DB 19; length 438;
XX Best Local Similarity 65.8%; Pred. NO. 1.4e-88;
XX Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
XX
QY 8 PCSRG-SSWSADLDKMCDCAS-----CRARP-----HSDFCLG- 39
DB 62 pctgfyneavnydtkqctqcnhrsgselkqncptqdtvcrcrptqrgsgyklgv 121
QY 40 -CAAAPAPAP-----RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDT 52
DB 122 dcvcpgphspgnnqackpwtncitlgskqtrhpasdsldavcedslatlilwetqrpt 181
QY 53 -----RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDT 85
DB 182 frptvtgwtvprtselstptltveprscdkthtccpcpapeagapsvflfppkpkdt 241
QY 86 LMTSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 145
DB 242 lmsrtpcvvvdvshedpevkfnwvvgvvevhnaktprceeqynstyrvvsvltvln 301
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVK 205
DB 302 qdwlngkeykckvsnkalpapiektiskakgqpreqvyltppsrdeltnqysltclvk 361
QY 206 GFPPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVSCSVMHE 265
DB 362 gfypsdiavewesngqpennnykttppvldsdgsfflyslkltvdksrwqggnvfscvmhe 421
QY 266 ALHNHYTQKSLSLSPGK 282
DB 422 alnhhytcqkslsispgk 438
XX
RESULT 5
AAV13463
ID AAY13463 standard; Protein; 660 AA.
XX
XX AAY13463;
XX
XX 26-JUL-1999 (first entry)
XX
XX Ectromelia A39R semaphorin polypeptide.
XX
XX VESPR; viral-encoded semaphorin protein receptor; semaphorin; IL-12;
XX inflammatory disease; viral infection; immune regulation; interleukin-12;
XX Th1 type immune response; cancer; autoimmune disease; A39.
XX
XX Ectromelia virus.
XX
XX WO921997-A1.
XX
```

```
PD 06-MAY-1999.
XX
XX 28-OCT-1998; 98WO-US22879.
XX
XX 28-OCT-1997; 97US-0112009.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX Comeau MR, Dubose RF, Johnson RS, Spriggs MK;
XX
XX WPI; 1999-326704/27.
XX N-PSDB; AAX55565.
XX
XX Viral-encoded protein receptors useful for treating inflammatory
XX diseases
XX
XX Example 1; Page 63-66; 73pp; English.
XX
XX The invention relates to a VESPR (viral-encoded semaphorin protein
XX receptor) polypeptide that can bind semaphorins. The VESPR polypeptides
XX are used to treat inflammatory diseases; when immobilized, to purify and
XX separate semaphorins or cells that express them. They are useful as
XX reagents for detecting, or measuring biological activity of, optionally
XX modified semaphorins; as carriers for delivering diagnostic and
XX therapeutic agents to semaphorin-positive cells; for studying the role of
XX VESPR in viral infection and immune regulation. When administered, in
XX combination with vaccination, together with a semaphorin, VESPR may
XX stimulate production of interleukin-12 (IL-12) which acts as an adjuvant
XX to induce a more persistent cellular immune response (of Th1 type),
XX including a curative response against aggressive, micrometastatic
XX cancers. Antibodies against VESPR can be used to treat autoimmune
XX diseases where an inflammatory response follows presentation of self-
XX antigens by T cells. Sense and antisense fragments of nucleic acid
XX encoding VESPR can be used to block expression of the polypeptide. The
XX present sequence represents an eromelia A39 semaphorin protein. The
XX native VESPR polypeptide is isolated from the human cells expressing the
XX receptor by using an Ectromelia virus A39 semaphorin/fc Fusion protein.
XX
XX Sequence 660 AA;
XX
XX Query Match 80.6%; Score 1250; DB 20; Length 660;
XX Best Local Similarity 100.0%; Pred. NO. 5.8e-88;
XX Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 53 RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDTLMTSRTPEVTCVVVDVSHEDPEVKFNW 112
DB 34 rscdkthtccpcpapeagapsvflfppkpkdtlmsrtpevtcvvvdvshedpevkfnw 93
QY 113 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
DB 94 yvdgvevhnaktprceeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 232
DB 154 kakgqpreqvyltppsreemtqnqvsaltclvkgyfypsdiavewesngqpennnykttppv 213
QY 233 LDSGSGFFLYSLKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPGK 282
DB 214 ldsdgsfflyslkltvdksrwqggnvfscsvmhealnhhytcqkslsispgk 263
XX
RESULT 6
AAB28523
ID AAB28523 standard; Protein; 660 AA.
XX
XX AAB28523;
XX
XX 07-FEB-2001 (first entry)
XX
XX Ectromelia A39R semaphorin.
XX
XX Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;
```



KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;  
KW inflammation; immune regulation; viral infection.  
XX Ectromelia virus.  
PN US6130068-A.  
XX 10-OCT-2000.  
XX 28-OCT-1998; 98US-0181706.  
XX 26-OCT-1998; 98US-0112009.  
XX (IMMV ) IMMUNEX CORP.  
PI Comeau MR, Johnson RS, Spriggs MK, Dubose RF;  
XX WPI; 2000-646753/62.  
DR N-PSDB; AAC63726.  
XX New Viral Encoded Semaphorin Protein Receptor DNA and polypeptides,  
PT useful for treating inflammation or inflammatory diseases, e.g.  
PT rheumatoid arthritis -  
XX Example 1; Column 55-60; 32pp; English.  
XX The present sequence is provided in an invention relating to novel  
CC semaphorin receptor polypeptides designated Viral Encoded Semaphorin  
CC Protein Receptor (VESPR). The DNA encoding the VESPR is useful for  
CC constructing expression vectors that produce the VESPR polypeptide. The  
CC VESPR DNA and polypeptide are useful for treating inflammation and  
CC inflammatory diseases, e.g. rheumatoid arthritis, a disease associated  
CC with chronic inflammation of the synovial tissue. The VESPR polypeptide  
CC is useful for measuring the biological activity of semaphorin proteins  
CC in terms of their binding affinity for VESPR. It is also useful as a  
CC research tool for studying the role that the receptor, in conjunction  
CC with semaphorins, may play in immune regulation and viral infection. The  
CC VESPR polypeptides are also useful as reagents that may be employed by  
CC those conducting quality assurance studies, e.g. to monitor shelf life  
CC and stability of semaphorin protein under different conditions.  
XX  
SQ Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 21; Length 660;  
Best Local Similarity 100.0%; Pred. No. 5.8e-88;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112  
DB 34 RSCDKTHTCPPCPAPEAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 113 YVDGVEVHNATKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172  
DB 94 YVDGVEVHNATKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153  
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 232  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213  
QY 233 LDSGDSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNYHTQKSLSISPGK 282  
DB 214 LDSGDSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNYHTQKSLSISPGK 263

RESULT 7  
AAB70132  
ID AAB70132 standard; Protein; 660 AA.  
XX  
AC AAB70132;  
XX  
DT 21-MAY-2001 (first entry)  
XX

DE Ectromelia A39R semaphorin.  
XX  
KW Vaccinia virus; Ectromelia; A39R; semaphorin; VESPR;  
KW viral encoded semaphorin protein receptor; antiinflammatory;  
KW inflammation.  
XX  
OS Vaccinia virus.  
PN US6187909-B1.  
XX 13-FEB-2001.  
PD 10-DEC-1999; 99US-0459066.  
PF 26-OCT-1998; 98US-0112009.  
PR 28-OCT-1998; 98US-0181706.  
PR 28-OCT-1997; 97US-0958598.  
XX (IMMV ) IMMUNEX CORP.  
XX  
XX Spriggs MK, Comeau MR, Dubose RF, Johnson RS;  
PI WPI; 2001-256141/26.  
XX N-PSDB; AAF76950.  
DR  
XX  
XX New viral encoded semaphorin protein receptor polypeptides useful for  
PT treating inflammation or diseases associated with pro-inflammatory  
PT activity of a semaphorin ligand -  
XX  
PS Example 1; Column 55-60; 32pp; English.  
XX  
XX The DNA encoding the present sequence was used to prepare an Ectromelia  
CC Semaphorin/Fc fusion protein. The fusion protein was used in the  
CC isolation of a human viral encoded semaphorin protein receptor (VESPR).  
CC The VESPR polypeptides are useful for treating inflammation, or diseases  
CC associated with pro-inflammatory activity of a semaphorin ligand. The  
CC polypeptides are also useful as carriers for delivering diagnostic or  
CC therapeutic agents to cells expressing semaphorins.  
XX  
SQ Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 22; Length 660;  
Best Local Similarity 100.0%; Pred. No. 5.8e-88;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112  
DB 34 RSCDKTHTCPPCPAPEAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 93  
QY 113 YVDGVEVHNATKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172  
DB 94 YVDGVEVHNATKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153  
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 232  
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213  
QY 233 LDSGDSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNYHTQKSLSISPGK 282  
DB 214 LDSGDSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNYHTQKSLSISPGK 263

RESULT 8  
AAB31694  
ID AAB31694 standard; Protein; 660 AA.  
XX  
AC AAB31694;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Amino acid sequence of the Ectromelia virus A39R semaphorin.  
XX

KW Viral encoded semaphorin protein receptor; VESPR; semaphorin; interferon;  
KW viral semaphorin ligand; Ectromelia A39R; Staphylococcus aureus; IL-12;  
KW interleukin-12; dendritic cell; cytokine production; immunomodulator;  
KW proinflammatory cytokine; Th1 cell differentiation; inflammation;  
KW inflammatory disease; tumour.

XX Ectromelia virus.

PN US6174689-B1.

XX 16-JAN-2001.

XX 10-DEC-1999; 99US-0458791.

XX 26-OCT-1998; 98US-0112009.

PR 28-OCT-1998; 98US-0181706.

PR 27-OCT-1997; 97US-0958598.

XX (IMMV ) IMMUNEX CORP.

XX Spriggs MK, Comeau MR, Dubose RF, Johnson RS;  
PI WPI; 2001-167789/17.

XX N-PSDB; AAF25254.

XX Screening for binding to viral encoded semaphorin protein receptor,  
PT comprises contacting a mixture containing semaphorin, or cells that  
PT express semaphorin with protein, and detecting binding to protein -  
XX Example 1; Column 55-60; 32pp; English.

XX The present sequenc represents an Ectromelia virus A39R semaphorin  
CC ligand. The protein was used to isolate a viral encoded semaphorin  
CC protein receptor (VESPR) polypeptide from human cells. Semaphorins  
CC interact with their membrane bound receptors to synergise with interferon  
CC and Staphylococcus aureus (type C) in the production of interleukin-12  
CC (IL-12) from dendritic cells. VESPR can therefore be used to induce  
CC IL-12 production which in turn promotes cytokine production. IL-12 is  
CC a proinflammatory cytokine and an immunomodulator. A soluble VESPR can  
CC be used to antagonise IL-12 and downregulate Th1 cell differentiation.  
CC VESPR is therefore useful for treating inflammation and inflammatory  
CC diseases. VESPR ligands may be used to induce an immune response against  
CC aggressive tumours.

XX Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 22; Length 660;  
Best Local Similarity 100.0%; Pred. No. 5.8e-88;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 112

Db 34 rscdkthtccpapeagapsvflppkpkdtlmisrtpevtcvvvdvshedpevkfnw 93

QY 113 YVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKREYCKVSKNKPAPIEKTIS 172

Db 94 yvdgvevhnaktkpreeqynstyrvvsvltvlgdwlngkeyckvsknkalpapiektis 153

QY 173 KAKQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 232

Db 154 kagqprepqvtytlppseemtqnqslctclvkgyfypsdiavewsgngqpennnykttppv 213

QY 233 LDSGSRFLYSKLTVDKSRWQQGNVFCGVMHEALHNHYTKLSLSPGK 282

Db 214 ldsdgsrflfyskltvdksrwqgnvfscgvmhealhnhytqklslspgk 263

RESULT 9

AAAY5239

XX AAY15239 standard; Protein: 450 AA.

XX AAY15239;

AC "

XX

DT 04-NOV-1999 (first entry)

DE ULBP1-Fc fusion polypeptide amino acid sequence.

XX cell surface glycoprotein; lymphoma; assay; retinitis pigmentosa;  
KW diabetes mellitus; progressive pseudo rheumatoid arthropathy;  
KW muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.  
KW fusion polypeptide.

XX Homo sapiens.

OS Synthetic.

XX WO9931241-A1.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-US27048.

XX 15-JUL-1998; 98US-0092946.

PR 17-DEC-1997; 97US-0069857.

XX (IMMV ) IMMUNEX CORP.

XX Cosman DJ, Fanslow WC, Mullberg JH;  
PI WPI; 1999-493777/41.

XX DNA encoding human B cell lymphocyte cell surface glycoproteins  
PT Disclosure; Page 29; 104pp; English.

XX This is the amino acid sequence of the ULBP1-Fc fusion polypeptide. The  
CC ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins  
CC using an Fc muclen to provide fused polypeptides, such as ULBP1-Fc.  
CC The polypeptides ULBP-1 and -2 bind UL16-Fc. In addition, they bind to a  
CC number of human cell types, including mitogen-stimulated human T cells  
CC and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an  
CC anaplastic lymphoma. The ULBP proteins can therefore be used as markers  
CC to detect cancer, to enhance IFN- gamma production, NK cell proliferation  
CC and CTL activity, to purify proteins and measure their activity. The  
CC polypeptides and their fragments can also be used as delivery and  
CC therapeutic agents, for rational drug design, as research reagents,  
CC controls for peptide fragmentation, molecular weight/isoelectric focusing  
CC markers, identification of unknown proteins and also for preparation of  
CC antibodies.

CC The antibodies can be used in assays to detect the presence of ULBP  
CC proteins, in vitro or in vivo, as well as for use in purification of ULBP  
CC proteins. The ULBP DNA sequences can be used as probes to identify  
CC homologues, to identify human chromosome number 6, to map and identify  
CC genes, especially associated with certain diseases, syndromes or other  
CC conditions on chromosome 6, as single-stranded sense or antisense  
CC oligonucleotides, to inhibit expression of ULBP polypeptides, to help  
CC detect defective genes in an individual and for gene therapy.  
CC Diseases, syndromes and conditions associated with human chromosome  
CC 6 include Retinitis pigmentosa (6q14-q21), Diabetes mellitus  
CC (insulin-dependent) (6q21), progressive pseudorheumatoid arthropathy of  
CC childhood (6q22), muscular dystrophy (congenital merosin-deficient)  
CC (6q22-q23), and cardiomyopathy (dilated) (6q23).

XX Sequence 450 AA;

Query Match 80.4%; Score 1247; DB 20; Length 450;  
Best Local Similarity 96.3%; Pred. No. 6.2e-88;  
Matches 232; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 44 PP--APFRLWRSCDKTHTCPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVDV 101

Db 210 ppslapgttprscdkthtccpapeagapsvflppkpkdtlmisrtpevtcvvvdv 269

QY 102 SHEDPEYKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNKREYCKVSKN 161

Db 270 shedpevfknvydvgevhnaaktppreeqynstyrsvvltvlhgdwlngkeyckvsnk 329

QY 162 ALPAPEKTIKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEENSGQ 221

Db 330 alpapiektiskakgqprepqvtytlppsrdeitknqsltcclvklgfydsdiavevesnqg 389

QY 222 PENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 281

Db 390 pennykttppvldsdgsfflyskltdvksrwqgnvfscsvmhealhnhytqkslspsg 449

QY 282 K 282

Db 450 k 450

RESULT 10

AAV15240

ID AAV15240 standard; Protein: 453 AA.

AC AAV15240;

XX

DT 04-NOV-1999 (first entry)

XX

DE ULBP2-Fc fusion polypeptide amino acid sequence.

XX

KW cell surface glycoprotein; lymphoma; assay: retinitis pigmentosa;

KW diabetes mellitus; progressive pseudo rheumatoid arthropathy;

KW muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.

KW fusion polypeptide.

XX

OS Homo sapiens.

OS

PN WO9931241-A1.

XX

PD 24-JUN-1999.

XX

PF 17-DEC-1998; 98WO-US27048.

XX

PR 15-JUL-1998; 98US-0092946.

PR 17-DEC-1997; 97US-0069857.

XX

PA (IMMV ) IMMUNEX CORP.

XX

XX Cosman DJ, Fanslow WC, Mullberg JH;

XX

DR WPI; 1999-493777/41.

XX

PT DNA encoding human B cell lymphocyte cell surface glycoproteins

XX

PS Disclosure; Page 29; 104pp; English.

XX

CC This is the amino acid sequence of the ULBP2-Fc fusion polypeptide. The

CC ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins

CC using an Fc mutain to provide fused polypeptides, such as ULBP2-Fc.

CC The polypeptides ULBP-1 and -2 bind UL16-Fc. In addition, they bind to a

CC number of human cell types, including mitogen-stimulated human T cells

CC and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an

CC anaplastic lymphoma. The ULBP proteins can therefore be used as markers

CC to detect cancer, to enhance IFN- gamma production, NK cell proliferation

CC and CTL activity, to purify proteins and measure their activity. The

CC polypeptides and their fragments can also be used as delivery and

CC therapeutic agents, for rational drug design, as research reagents,

CC controls for peptide fragmentation, molecular weight/isoelectric focusing

CC markers, identification of unknown proteins and also for preparation of

CC antibodies.

CC The antibodies can be used in assays to detect the presence of ULBP

CC proteins, in vitro or in vivo, as well as for use in purification of ULBP

CC proteins. The ULBP DNA sequences can be used as probes to identify

CC homologues, to identify human chromosome number 6, to map and identify

CC genes, especially associated with certain diseases, syndromes or other

CC conditions on chromosome 6, as single-stranded sense or antisense

CC oligonucleotides, to inhibit expression of ULBP polypeptides, to help

CC detect defective genes in an individual and for gene therapy.

CC Diseases, syndromes and conditions associated with human chromosome

CC 6 include Retinitis pigmentosa (6q14-q21), Diabetes mellitus

CC (insulin-dependent) (6q21), progressive pseudorheumatoid arthropathy of

CC childhood (6q22), Muscular dystrophy (congenital merosin-deficient)

CC (6q22-q23), and cardiomyopathy (dilated) (6q23).

XX

SQ Sequence 453 AA;

Query Match 80.4%; Score 1247; DB 20; Length 453;

Best Local Similarity 91.5%; Pred. No. 6.2e-88;

Matches 236; Conservative 3; Mismatches 9; Indels 10; Gaps 2;

QY 35 DFCLG-----CAAAPPA---PFRLLWRSCDKTHTCPCPAPEAGAPSVFLFPKPKD 84

Db 196 dfllmgdstlepsagaplamssgtqlrrscdkhtcpcpapeagapsvflfpkpkd 255

QY 85 TLMISRTPEVTCVVVDVSHEDPEVKFNQYVGVGVHNNAKTKPREEQYNSTYRVVSVLTVL 144

Db 256 tlmisrtpevtcvcvvdvshedpevkfnvydvgevhnaaktppreeqynstyrsvvltvl 315

QY 145 HQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLV 204

Db 316 hqdwlngkeyckvsnkalpapiektiskakgqprepqvtytlppsrdeitknqsltciv 375

QY 205 KGFYPSDIAVEENSGQPNENYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMH 264

Db 376 kgyfypsdiavevesngqpennyykttppvldsdgsfflyskltdvksrwqgnvfscsvmh 435

QY 265 EALHNHYTKQSLSPGK 282

Db 436 ealhnhytqkslspsgk 453

RESULT 11

AAW14764

ID AAW14764 standard; Protein: 424 AA.

XX

AC AAW14764;

XX

DT 11-JUN-1997 (first entry)

XX

DE Human soluble kit ligand-IgG fusion protein.

XX

KW Kit ligand; c-kit proto-oncogene; cytokine; growth factor;

KW haematopoietic cell; cell proliferation; stem cell; anaemia;

KW thrombocytopaenia; therapy; IgG1.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..25

FT /label= Sig\_peptide

FT /note= "KL signal peptide"

FT Protein 26..424

FT /label= Mat\_protein

FT /note= "human KL-IgG fusion"

XX

W09526199-A1.

XX

PD 05-OCT-1995.

XX

PF 28-MAR-1995; 95WO-US03866.

XX

PR 28-MAR-1994; 94US-0220379.

XX

PA (CYTO-) CYTOMED INC.

XX

PI Lobell RB, Nocka KH;

XX

DR WPI; 1995-351198/45.

DR N-PSDB; AAT63109.

XX

PT Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit  
 PT increased activity in promoting cell proliferation  
 XX  
 XX  
 PS Claim 10; Page 43-44; 88pp; English.

XX A fusion protein (AAW14764) between human soluble kit ligand (KL)  
 CC (see also AAW14761) and a human IgG1 heavy chain can be transiently  
 CC expressed in COS cells transfected with a human KL-Ig cDNA  
 CC construct (AAT63109) in vector CDM8; a corrected KL-Ig construct  
 CC (AAW14765) has also been prep'd. KL-Ig can also be produced as a  
 CC dimer stabilised by intermolecular disulphide bonds or a peptide  
 CC linker. The stabilised KL-Ig dimers have a more favorable cell  
 CC proliferation: mast cell activation ratio than native KL and can  
 CC stimulate haematopoietic recovery or stem cell/progenitor cell  
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 80.1%; Score 1243; DB 16; Length 424;  
 Best Local Similarity 84.6%; Pred. No. 1.2e-87;  
 Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;

QY 3 APGTAPCSGRSSWSADLDKCMDCASCARPHSDFCGCAAPAPFRLLRSCDKTHTCP 62  
 Db 157 asetsdcvssstlspekds-----rvsvtkpflmfpvaadpep-----kscdkthtcp 204  
 QY 63 PCPAPEAEGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 122  
 Db 205 pcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwvdgvevhn 264  
 QY 123 KTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 182  
 Db 265 ktkpreeqynstyrvvsvltvllhqdwlngkeyckvsnkalpapiektiskakgqprepq 324  
 QY 183 VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 242  
 Db 325 vytlppsrdeitknqvsitclvkgfypsdiavewesngqpennykttppvldsdgsfely 384  
 QY 243 SKLTVDSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK 282  
 Db 385 skltvdksrwqggnvfscsvmhealhnhytqkslsispgk 424

RESULT 12

AAW14765  
 ID AAW14765 standard; Protein: 424 AA.

XX AAW14765;

XX 11-JUN-1997 (first entry)

XX Human soluble kit ligand-IgG fusion protein (corrected).

XX Kit ligand; c-kit proto-oncogene; cytokine; growth factor;  
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;  
 KW thrombocytopaenia; therapy; IgG1.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..25  
 FT /label= Sig\_peptide  
 FT /note= "KL signal peptide"  
 FT Protein 26..424  
 FT /label= Mat\_protein  
 FT /note= "human KL-Ig fusion"

PN W09526199-A1.

XX 05-OCT-1995.

XX 28-MAR-1995; 95WO-US03866.

XX 28-MAR-1994; 94US-0220379.  
 XX (CYTO-) CYTOMED INC.  
 XX  
 XX Lobell RB, Nocka KH;  
 XX WPI: 1995-351198/45.  
 DR N-PSDB; AAT63110.

XX Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit  
 PT increased activity in promoting cell proliferation  
 XX  
 XX Claim 10; Page 46-48; 88pp; English.

XX A fusion protein (AAW14765) between human soluble kit ligand (KL)  
 CC (see also AAW14761) and a human IgG1 heavy chain can be transiently  
 CC expressed in COS cells transfected with a human KL-Ig cDNA  
 CC construct (AAT63110) in vector CDM8. KL-Ig can also be produced as a  
 CC dimer stabilised by intermolecular disulphide bonds or a peptide  
 CC linker. The stabilised KL-Ig dimers have a more favorable cell  
 CC proliferation: mast cell activation ratio than native KL and can  
 CC stimulate haematopoietic recovery or stem cell/progenitor cell  
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 80.1%; Score 1243; DB 16; Length 424;  
 Best Local Similarity 84.6%; Pred. No. 1.2e-87;  
 Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;

QY 3 APGTAPCSGRSSWSADLDKCMDCASCARPHSDFCGCAAPAPFRLLRSCDKTHTCP 62  
 Db 157 asetsdcvssstlspekds-----rvsvtkpflmfpvaadpep-----kscdkthtcp 204  
 QY 63 PCPAPEAEGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 122  
 Db 205 pcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwvdgvevhn 264  
 QY 123 KTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 182  
 Db 265 ktkpreeqynstyrvvsvltvllhqdwlngkeyckvsnkalpapiektiskakgqprepq 324  
 QY 183 VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 242  
 Db 325 vytlppsrdeitknqvsitclvkgfypsdiavewesngqpennykttppvldsdgsfely 384  
 QY 243 SKLTVDSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK 282  
 Db 385 skltvdksrwqggnvfscsvmhealhnhytqkslsispgk 424

RESULT 13

AAW60037  
 ID AAW60037 standard; Protein: 376 AA.

XX AAW60037;

XX 11-SEP-1998 (first entry)

XX Antigenic peptide hFas (nd29) containing FC region.

XX Fas ligand; Fas antagonist; apoptosis related disease; liver disease;  
 KW heart failure; kidney failure; graft-versus-host disease; antibody;  
 KW myocardial infarction; ischemic restenosis; endotoxic shock.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..16  
 FT /note= "hFas antigen signal peptide"  
 FT Protein 30..376



CC polypeptides are proteinases implicated in fertilisation and  
CC spermatogenesis. They can be used as therapeutic agents. A proteinase  
CC inhibitor of the catalytic domain would inhibit SVPH1-26 activity and  
CC would be useful as a method for birth control. Also, an inhibitor of the  
CC disintegrin domain of SVPH1-26 may affect fertilisation. The proteinase  
CC activity of SVPH1-26 can also be used as a detergent additive for the  
CC removal of stains having a protein component. The SVPH1-26 polypeptides  
CC and fragments can also be used as molecular weight markers, as markers  
CC for determination of isoelectric points of sample proteins and as  
CC controls for establishing the extent of fragmentation of a protein  
CC sample. The products can also be used for identifying, separating or  
CC purifying cells that express SVPH1-26 polypeptides such as testis cells.  
CC They can be used for the diagnosis and prognosis of testicular cancers.  
CC The present sequence represents a fusion protein from an example of the  
CC present invention comprising: the Ig kappa signal sequence at positions  
CC 1 to 20; a spacer at positions 21 to 22; the SVPH1-26 disintegrin domain  
CC at positions 23 to 305; a spacer at positions 306 to 307; and the  
CC IgG1 Fc domain at positions 308 to 535.  
XX Sequence 535 AA;  
SQ

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:13:16 ; Search time 54.85 Seconds  
(without alignments)  
824.032 Million cell updates/sec

Title: US-09-742-454A-7  
Perfect score: 1678  
Sequence: 1 MARGSLRLRLRLVLGLWLA.....MHEALHNHYTKQSLSLSPCK 309

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID     | Description         |
|------------|-------|---------------|--------|-----------|---------------------|
| 1          | 843   | 50.2          | 463    | 11 Q99LC4 | Q99LC4 mus musculus |
| 2          | 839   | 50.0          | 437    | 11 Q9RLA4 | Q9RLA4 mus musculus |
| 3          | 804.5 | 47.9          | 473    | 11 Q9D8L4 | Q9D8L4 mus musculus |
| 4          | 799   | 47.6          | 473    | 11 Q99L25 | Q99L25 mus musculus |
| 5          | 795.5 | 47.4          | 468    | 11 Q99L31 | Q99L31 mus musculus |
| 6          | 431.5 | 25.7          | 129    | 4 Q9NP84  | Q9NP84 homo sapien  |
| 7          | 367   | 21.9          | 375    | 4 Q9BS21  | Q9BS21 homo sapien  |
| 8          | 367   | 21.9          | 597    | 4 Q9BU10  | Q9BU10 homo sapien  |
| 9          | 367   | 21.9          | 597    | 4 Q9BOB8  | Q9BOB8 homo sapien  |
| 10         | 335.5 | 20.0          | 129    | 11 Q9QZK3 | Q9QZK3 mus musculus |
| 11         | 330.5 | 19.7          | 129    | 11 Q9CR75 | Q9CR75 mus musculus |
| 12         | 284.5 | 17.0          | 384    | 4 Q9UP60  | Q9UP60 homo sapien  |
| 13         | 279.5 | 16.7          | 500    | 4 Q9BRV0  | Q9BRV0 homo sapien  |
| 14         | 274.5 | 16.4          | 416    | 4 Q9NP66  | Q9NP66 homo sapien  |
| 15         | 270   | 16.1          | 487    | 11 Q99KA4 | Q99KA4 mus musculus |
| 16         | 269.5 | 16.1          | 684    | 13 Q90544 | Q90544 ginglymosto  |
| 17         | 269   | 16.0          | 426    | 11 Q9DCD9 | Q9DCD9 mus musculus |
| 18         | 255   | 15.2          | 479    | 11 Q99M22 | Q99M22 mus musculus |
| 19         | 255   | 15.2          | 484    | 11 Q99LA6 | Q99LA6 mus musculus |

|    |       |      |     |           |                     |
|----|-------|------|-----|-----------|---------------------|
| 20 | 197   | 11.7 | 94  | 4 Q9HCS0  | Q9HCS0 homo sapien  |
| 21 | 185   | 11.0 | 268 | 13 Q90524 | Q90524 ginglymosto  |
| 22 | 184.5 | 11.0 | 130 | 11 Q9D8W4 | Q9D8W4 mus musculus |
| 23 | 184   | 11.0 | 509 | 11 Q9QX57 | Q9QX57 mus musculus |
| 24 | 184   | 11.0 | 513 | 11 P97797 | P97797 mus musculus |
| 25 | 182.5 | 10.9 | 509 | 11 Q9WTN4 | Q9WTN4 mus musculus |
| 26 | 181   | 10.8 | 235 | 11 Q99M11 | Q99M11 mus musculus |
| 27 | 176.5 | 10.5 | 259 | 13 Q90530 | Q90530 ginglymosto  |
| 28 | 176   | 10.5 | 506 | 6 Q46632  | Q46632 bos taurus   |
| 29 | 173   | 10.3 | 509 | 11 Q08907 | Q08907 mus musculus |
| 30 | 172   | 10.3 | 506 | 6 Q46631  | Q46631 bos taurus   |
| 31 | 171.5 | 10.2 | 257 | 13 Q90536 | Q90536 ginglymosto  |
| 32 | 170   | 10.1 | 237 | 13 Q90545 | Q90545 ginglymosto  |
| 33 | 168   | 10.0 | 105 | 11 Q99JC1 | Q99JC1 mus musculus |
| 34 | 167.5 | 10.0 | 252 | 13 Q90568 | Q90568 ginglymosto  |
| 35 | 166.5 | 9.9  | 267 | 13 Q90529 | Q90529 ginglymosto  |
| 36 | 165   | 9.8  | 238 | 7 Q9MXA2  | Q9MXA2 aulonocara   |
| 37 | 162   | 9.7  | 261 | 7 Q19363  | Q19363 sus scrofa   |
| 38 | 160   | 9.5  | 261 | 7 Q62868  | Q62868 sus scrofa   |
| 39 | 157.5 | 9.4  | 237 | 7 Q9MX99  | Q9MX99 aulonocara   |
| 40 | 155   | 9.2  | 261 | 7 Q98263  | Q98263 sus scrofa   |
| 41 | 152.5 | 9.1  | 354 | 4 Q9NQR8  | Q9NQR8 homo sapien  |
| 42 | 151   | 9.0  | 238 | 11 Q99X37 | Q99X37 mus musculus |
| 43 | 150   | 8.9  | 208 | 7 Q9MXA0  | Q9MXA0 aulonocara   |
| 44 | 150   | 8.9  | 260 | 7 P79551  | P79551 homo sapien  |
| 45 | 149.5 | 8.9  | 354 | 4 Q9P1W8  | Q9P1W8 homo sapien  |

## ALIGNMENTS

RESULT 1  
Q99LC4 ID AC Q99LC4; PRELIMINARY; PRT; 463 AA.  
AD Q99LC4;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003435; AAH03435.1; -  
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 50.2%; Score 843; DB 11; Length 463;  
Best Local Similarity 46.8%; Pred. No. 2.3e-69;  
Matches 161; Conservative 45; Mismatches 62; Indels 76; Gaps 6;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 24  | SVAGEQAPGTAPCSRGSWSADLDKCMDCASCACRAPHSDCLGCAAP--PAPFRLWRS   | 81  |
| Db | 138 | SAAKTTPPSVYPLAPCSA-----AOTNSMTLCLVKGYFPEPVVTWNS             | 182 |
| Qy | 82  | -----CDKTHT-----  | 87  |
| Db | 183 | GSLSGVTFFPAVLQSLDLYTLSSSVTPSPSTWPTVCNVAHPASSTKVKDKKIVPRDCG  | 242 |
| Qy | 88  | CPP--CPAPEARGASVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE  | 145 |
| Db | 243 | CKPKICTTPEVS---SVETFPKPKDVLITITLTPKVTCTVVVDISKDDPEVQSFVDDVD | 299 |
| Qy | 146 | VHNAKTPREEQYNSTYRVSVLTIVLHQLNGKVEKCKVSNKALPAPTEKTSRAKGOP    | 205 |
| Db | 300 | VHTAQTPREEQFNSTFRSVSELPIMHQDLNGKFKCKRVNSAAPPAPTEKTSIKTKGRP  | 359 |
| Qy | 206 | REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS | 265 |

```
Db 360 KAPQVYTIPTPPKQMAKDQVSLTCMTIDFFPEDITVWQWNGQPAENYKNTQPIMDTGDG 419
QY 266 FFYLSKLTVDKSWQGNVFCSCVMHEALHNNHYTKQSLSPCK 309
Db 420 YFYISKLNVQKSWNEAGNTFTCSVLHGLHNNHTEKSLSPCK 463

RESULT 2
Q9R1A4
ID Q9R1A4 PRELIMINARY; PRT; 437 AA.
AC Q9R1A4;
DT 01-MAY-2000 (TREMELREL. 13, Created)
DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)
DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)
DE CAMVAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO IMMNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 50.0%; Score 839; DB 11; Length 437;
Best Local Similarity 46.5%; Pred. No. 5e-69;
Matches 160; Conservative 46; Mismatches 62; Indels 76; Gaps 6;

QY 24 SVAGEQAPGTAPCSRGSWSADLDKCMDCASCRCRPHSDFCLGCAAP---PAPFRLLRWS 81
Db 112 SAAKTTPSVYPLAGSA-----AQTNSMVTGLGVKGYPPPTVTWNS 156
QY 82 -----CDKTH----- 87

Db 157 GLSGSGVHTFPAYLQSDLYTLSSSVTPSPSTVTCNVAHPASSTKVDKXIVPRDCG 216
QY 88 CPP--CPAPEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGE 145
Db 217 CKFCICTVPEVS---SVFIFPPKPKDVLITITLPKTCVVVDISKDDPEVQFSWFVDDE 273
QY 146 VHNAKTKPREQNSTYRVVSVLTVLHQDLNGLNGKEYCKVSKNALKALPAPIEKTISKAGQP 205
Db 274 VHTAQTPREQNSTFRSVSELPIMHQDLNGLNGKEYCKVSKNAAFPAPIEKTISKGRP 333
QY 206 REPOVTLPSREEMTKNQVSLTCLVKGYFYPVDIAVWESNGQPENNYKTPPVLDSDGS 265
Db 334 KAPQVYTIPTPPKQMAKDQVSLTCMTIDFFPEDITVWQWNGQPAENYKNTQPIMDTGDG 393
QY 266 FFYLSKLTVDKSWQGNVFCSCVMHEALHNNHYTKQSLSPCK 309
Db 394 YFYISKLNVQKSWNEAGNTFTCSVLHGLHNNHTEKSLSPCK 437

RESULT 3
Q9D8L4
```

```
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4;
DT 01-JUN-2001 (TREMELREL. 17, Created)
DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)
DE 1810060009RIK PROTEIN.
GN 1810060009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: TO IMMNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AK007918; BAB25349.1; -.
DR MGD; MGI:1924014; 1810060009RIK.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; Ig_2.
DR SMART; SM00407; IGV; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 47.9%; Score 804.5; DB 11; Length 473;
Best Local Similarity 53.1%; Pred. No. 8.1e-66;
Matches 154; Conservative 44; Mismatches 83; Indels 9; Gaps 3;

QY 20 ALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRCRPHSDFCLGCAAPAPFRLRW 79
Db 193 ALLQS--GLYTLSSSVTVTNTWPSQTITCN-----VAHPASSTKVDKTEPRVP---ITQ 243
QY 80 RSCDKHTHTCPPEAEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
Db 244 NCPPLKECPCAAPDLLGGPSVFIFPKIKDVLMSLSPWTCVVVDVSHEDDQVLSW 303
QY 140 YDGVGEVHNAKTKPREQNSTYRVVSVLTVLHQDLNGLNGKEYCKVSKNALKALPAPIEKTIS 199
Db 304 FVNVVEVHTAQTHREDYNTLRVVSALPIQHDWMSGKEFKCKVNNRPLSPSIEKTIS 363
QY 200 KAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFYPVDIAVWESNGQPENNYKTPPV 259
Db 364 KPRGPVRAPQVYVLPPLPPPAEEMTKKESLTCMTITGFLPAEIAVDWTSNGRTEQNKNTATV 423
```



[illegible]

малышата, бусиетта, сатагитилт, номинтаде; номло:



```
QY 40 SWSADLDKCMDCASCARPHSDFCLGCAA---PPAPFRLWRS----- 81
Db 30 SSASPTSPKVFPLSLCSTQDGNVWVIAVLQVGFPOEPLSVTWSESQGVTAARFPSPSQD 89
QY 82 -----CDKTHTC-----PPCPAPEAGAPSVFLFPPK 108
Db 90 ASGDLYTTSSQLTLPATQCLAGKSVTHVKNYTPNSQDVTVPVSTPTTSPST-PPT 148
QY 109 PK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKP 153
Db 149 PSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGVFTTTPSPSGK--SAVQGP 205
QY 154 REEQNSTYRVVSVLVTHLDWLNKEYCKVSNKALPAPIETISKAKQPREPOVYTL 213
Db 206 PERDLGCGYSVSVLPCCAEFPWNGKFTCTAAYPESKTPLTATLSKS-GNTRFPEVHLL 264
QY 214 PPSREEMTKNQ-VSLTCLVKGYFSPDAVFNESNGQ--PENNYKTTTPPVLD-SDG--SFF 267
Db 265 PPPSEELALNELVTLTCLARGFSPKDVLRWLQGSQELPREKYLTVASRQEPSPSGTTTFA 324
QY 268 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 309
Db 325 VTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 366

RESULT 13
Q9BRV0 PRELIMINARY; PRT; 500 AA.
AC Q9BRV0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:14588).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005951; AAH05951.1; -.
SQ SEQUENCE 500 AA; 54154 MW; 0A9BF43F2A3CC6D9 CRC64;

Query Match 16.7%; Score 279.5; DB 4; Length 500;
Best Local Similarity 25.4%; Pred. No. 1.6e-17;
Matches 87; Conservative 41; Mismatches 137; Indels 77; Gaps 12;

QY 40 SWSADLDKCMDCASCARPHSDFCLGCAA---PPAPFRLWRS----- 81
Db 146 SSASPTSPKVFPLSLCSTQDGNVWVIAVLQVGFPOEPLSVTWSESQGVTAARFPSPSQD 205
QY 82 -----CDKTHTC-----PPCPAPEAGAPSVFLFPPK 108
Db 206 ASGDLYTTSSQLTLPATQCLAGKSVTHVKNYTPNSQDVTVPVSTPTTSPST-PPT 264
QY 109 PK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKP 153
Db 265 PSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGVFTTTPSPSGK--SAVQGP 321
QY 154 REEQNSTYRVVSVLVTHLDWLNKEYCKVSNKALPAPIETISKAKQPREPOVYTL 213
Db 322 PDRDLGCGYSVSVLPCCAEFPWNGKFTCTAAYPESKTPLTATLSKS-GNTRFPEVHLL 380
QY 214 PPSREEMTKNQ-VSLTCLVKGYFSPDAVFNESNGQ--PENNYKTTTPPVLD-SDG--SFF 267
Db 381 PPPSEELALNELVTLTCLARGFSPKDVLRWLQGSQELPREKYLTVASRQEPSPSGTTTFA 440
QY 268 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 309
```

```
Db 441 VTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 482

RESULT 14
Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RL "The European IMAGE consortium for integrated Molecular analysis of
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AL389978; CAB97534.1; -.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 16.4%; Score 274.5; DB 4; Length 416;
Best Local Similarity 29.6%; Pred. No. 3.8e-17;
Matches 81; Conservative 38; Mismatches 116; Indels 39; Gaps 12;

QY 48 KCMDC--CASCRRAR---PHSDFCLGCAAPAPFRLWRSCKDTHTCPCCPAPEAGCAPS 101
Db 152 QCPDGKSVTCHVKHYTNPSQDVTVPVPPP-----PPCHPR----- 189
QY 102 VFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKPREEQYNST 161
Db 190 LSLHRPALED-LLLGSEANLTCTLTGL-RDASGATFTTTPSSGK--SAVQGPPELDLGGC 245
QY 162 YRVSVLTVLHQDWLNKEYCKVSNKALPAPIETISKAKQPREPOVYTLPPSREMT 221
Db 246 YSVSVLPGLCAQPNWNGHGETTCTAAHPKLTPLTANITKS-GNTRFPEVHLLPPPSBELA 304
QY 222 KNO-VSLTCLVKGYFSPDAVFNESNGQ--PENNYKTTTPPVLD-SDG--SFELYSKLTVD 275
Db 305 LNELVTLTCARGFSPKDVLRWLQGSQELPREKYLTVASRQEPSPSGTTTFAVTSILRVA 364
QY 276 KSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 309
Db 365 AEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 398

RESULT 15
Q99KA4 PRELIMINARY; PRT; 487 AA.
AC Q99KA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE UNKNOWN (PROTEIN FOR MGC:6727).
```

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID-10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC004786; AAH04786.1; -  
SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

Query Match 16.1%; Score 270; DB 11; Length 487;  
Best Local Similarity 30.6%; Pred. No. 1.2e-16;  
Matches 77; Conservative 41; Mismatches 100; Indels 34; Gaps 11;

Qy 67 CAAPAPPAPFRLLRSCDKHTCPCPAPEAGAPSVFLFPKPKDITLMISRTPEVTCVV 126  
|: ||  
Db 243 CSGPPP-----PCPPCP-PSCH--PSLSLQRPALD-LLLGSDASLTCTLN 284  
|: ||  
Qy 127 DVSHEDPEVKFNWYVDGVEVHNATKPREQYNST---YRVSVLTVLHQDLNGKEYKC 183  
|: ||  
Db 285 GLRNPAGAV-FTW-----EPSTGKDAVQKKAQVNSCGCYSSVLPGCAERNWSGAFKC 338  
|: ||  
Qy 184 KVSNKALPAPIETKITSKAGQPREPOVYTLPPSREEMTKNO-VSLTCLVKGFPSPDIAVE 242  
|: ||  
Db 339 TVTHPESDT-LTGTLAKITVNTFFPQVHLPPPEELALNELVSLTCLVRAFNPKVLR 397  
|: ||  
Qy 243 WESNGQ---PENNYKTTTPPVLDSDG--SFFLYSKLTVDKSRWQGNVFCSYMHEALHNH 297  
|: ||  
Db 398 WLGNEELSPESYLVFEPKPEGEGATTYLVTSVLRVSAETWKQGDQYSCMVGHEALPMN 457  
|: ||  
Qy 298 YTKSLSLSPGK 309  
|: ||  
Db 458 FTQKTIDRLSGK 469

Search completed: March 11, 2002, 15:23:24  
Job time: 608 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:16 ; Search time 19.78 Seconds  
(without alignments)  
572.772 Million cell updates/sec

Title: US-09-742-454A-7

Perfect score: 1678

Sequence: 1 MARGSLRRLRLVLGLWLA.....MHEALHHVYTKSLSPGK 309

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description         |
|------------|--------|---------------|--------|-------|---------------------|
| 1          | 1224   | 72.9          | 330    | 1     | GC1_HUMAN           |
| 2          | 1159.5 | 69.1          | 326    | 1     | P01857 homo sapien  |
| 3          | 1151.5 | 68.6          | 290    | 1     | GC2_HUMAN           |
| 4          | 1139   | 67.9          | 327    | 1     | P01859 homo sapien  |
| 5          | 910.5  | 54.3          | 323    | 1     | GC4_HUMAN           |
| 6          | 876    | 52.2          | 329    | 1     | P01861 homo sapien  |
| 7          | 838.5  | 50.0          | 324    | 1     | P01870 oryctolagus  |
| 8          | 838    | 49.9          | 326    | 1     | P01862 cavia porce  |
| 9          | 836    | 49.8          | 329    | 1     | P01868 mus musculus |
| 10         | 833.5  | 49.7          | 333    | 1     | P020759 rattus norv |
| 11         | 825    | 49.2          | 333    | 1     | P22436 mus musculus |
| 12         | 825    | 49.2          | 398    | 1     | P01869 mus musculus |
| 13         | 817    | 48.7          | 329    | 1     | P20761 rattus norv  |
| 14         | 804.5  | 47.9          | 322    | 1     | P03987 mus musculus |
| 15         | 799.5  | 47.6          | 330    | 1     | P20762 rattus norv  |
| 16         | 794.5  | 47.3          | 335    | 1     | P01863 mus musculus |
| 17         | 794.5  | 47.3          | 399    | 1     | P01864 mus musculus |
| 18         | 787    | 46.9          | 336    | 1     | P01865 mus musculus |
| 19         | 782    | 46.6          | 405    | 1     | P01866 mus musculus |
| 20         | 367    | 21.9          | 454    | 1     | P01867 mus musculus |
| 21         | 366.5  | 21.8          | 421    | 1     | P01871 homo sapien  |
| 22         | 365    | 21.8          | 391    | 1     | P06336 mus musculus |
| 23         | 360.5  | 21.5          | 429    | 1     | P04220 homo sapien  |
| 24         | 353.5  | 21.1          | 428    | 1     | P01855 rattus norv  |
| 25         | 349    | 20.8          | 455    | 1     | P01854 homo sapien  |
| 26         | 344    | 20.5          | 458    | 1     | P01872 mus musculus |
| 27         | 339    | 20.2          | 476    | 1     | P03988 oryctolagus  |
| 28         | 336    | 20.0          | 454    | 1     | P01873 mus musculus |
| 29         | 334    | 19.9          | 479    | 1     | P06337 mesocricetu  |
| 30         | 333    | 19.8          | 457    | 1     | P04221 oryctolagus  |
| 31         | 331.5  | 19.8          | 450    | 1     | P20768 suncus muri  |
| 32         | 302.5  | 18.0          | 299    | 1     | P01874 canis fami   |
| 33         | 290    | 17.3          | 438    | 1     | P01879 oryctolagus  |
|            |        |               |        |       | P23085 heterodontu  |

|    |       |      |     |   |                     |
|----|-------|------|-----|---|---------------------|
| 34 | 288.5 | 17.2 | 446 | 1 | MUC_CHICK           |
| 35 | 284.5 | 17.0 | 438 | 1 | P23087 heterodontu  |
| 36 | 282   | 16.8 | 353 | 1 | P01876 homo sapien  |
| 37 | 279   | 16.6 | 353 | 1 | P20758 gorilla gor  |
| 38 | 275.5 | 16.4 | 461 | 1 | P23088 heterodontu  |
| 39 | 274.5 | 16.4 | 340 | 1 | P01877 homo sapien  |
| 40 | 274   | 16.3 | 370 | 1 | P23084 heterodontu  |
| 41 | 273.5 | 16.3 | 393 | 1 | P23086 heterodontu  |
| 42 | 255   | 15.2 | 344 | 1 | P01878 mus musculus |
| 43 | 252.5 | 15.0 | 481 | 1 | P23735 ictalurus p  |
| 44 | 178.5 | 10.6 | 105 | 1 | P01843 mus musculus |
| 45 | 173   | 10.3 | 104 | 1 | P20767 rattus norv  |

#### ALIGNMENTS

|  |   |
|--|---|
| RESULT 1   |   |
| GC1_HUMAN  | STANDARD; PRT; 330 AA.                        |
| AC P01857;   |   |
| DT 21-JUL-1986   | (Rel. 01, Created)                            |
| DT 21-JUL-1986   | (Rel. 01, Last sequence update)               |
| DT 15-JUL-1999   | (Rel. 38, Last annotation update)             |
| DE IG GAMMA-1  | CHAIN C REGION.                               |
| GN IGHG1.  |   |
| OS Homo sapiens  | (Human).                                      |
| OC Eukaryota; Metazoa;   | Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC Mammalia; Eutheria;   | Primates; Catarrhini; Hominiidae; Homo.       |
| OX NCBI_TaxID=9606;  |   |
| RP [1]   |   |
| RP SEQUENCE FROM N.A.  |   |
| RX MEDLINE=82274238;   | PubMed=6287432;                               |
| RA Ellison J.W., Berson B.J., Hood L.E.;                                 |   |
| RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene."    |   |
| RL Nucleic Acids Res. 10:4071-4079(1982).                                |   |
| RP [2]   |   |
| RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).                               |   |
| RX MEDLINE=71064024;   | PubMed=5489771;                               |
| RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,            |   |
| RA Waxdal M.J., Edelman G.M.;  |   |
| RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino |   |
| RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4."       |   |
| RL Biochemistry 9:3161-3170(1970).                                       |   |
| RP [3]   |   |
| RP SEQUENCE OF 136-329 (EU).   |   |
| RX MEDLINE=71064025;   | PubMed=5530842;                               |
| RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,         |   |
| RA Edelman G.M.;   |   |
| RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino   |   |
| RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7."       |   |
| RL Biochemistry 9:3171-3181(1970).                                       |   |
| RP [4]   |   |
| RP SEQUENCE (MYELOMA PROTEIN NIE).                                       |   |
| RX MEDLINE=77070269;   | PubMed=826475;                                |
| RA Poustingl H., Hilschmann N.;  |   |
| RT "The rule of antibody structure. The primary structure of a           |   |
| RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The        |   |
| RT chymotryptic peptides of the H-chain, alignment of the tryptic        |   |
| RT peptides and discussion of the complete structure."                   |   |
| RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).                 |   |
| RP [5]   |   |
| RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.                  |   |
| RX MEDLINE=83289131;   | PubMed=6884994;                               |
| RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;                     |   |
| RT "Three-dimensional structure determination of antibodies. Primary     |   |
| RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."     |   |
| RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).                   |   |
| RP [6]   |   |
| RP DISULFIDE BONDS.  |   |
| RX MEDLINE=71064027;   | PubMed=4923144;                               |
| RA Gall W.E., Edelman G.M.;  |   |
| RT "The covalent structure of a human gamma G-immunoglobulin. X.         |   |

Intrachain disulfide bonds.";  
Biochemistry 9:3188-3196(1970).  
[7]  
DISULFIDE BONDS.  
MEDLINE-77070267; PubMed-1002129;  
Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
"Rule of antibody structure. The primary structure of a monoclonal  
IgG1 immunoglobulin (myeloma protein Nie). I: Purification and  
characterization of the protein, the L- and H-chains, the  
cyanogen bromide cleavage products, and the disulfide bridges.";  
Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
[8]  
X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
MEDLINE-81208100; PubMed-7236608;  
Daisenhofer J.;  
"Crystallographic refinement and atomic models of a human Fc fragment  
and its complex with fragment B of protein A from Staphylococcus  
aureus at 2.9- and 2.8-A resolution.";  
Biochemistry 20:2361-2370(1981).  
CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
MARKER & THE GIM (NON-1) MARKERS.  
CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
35,116,198,269 & 272.  
CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
268-272.  
CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
RESIDUES 198,267&272.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; J00228; AAC82527.1; ALT\_INIT.  
PIR; A02146; GHU.  
DR PDB; 1FC1; 15-JUL-92.  
DR PDB; 1FC2; 15-JUL-92.  
DR MIM; 147100; -.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003597; Ig\_c1.  
DR InterPro; IPR003600; Ig\_like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; Igcl; 2.  
DR SMART; SM00410; Ig\_like; 1.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
3D-structure.  
FT NON\_TER 1 1  
FT DOMAIN 1 98 CH1.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 223 CH2.  
FT DOMAIN 224 330 CH3.  
FT DISULFID 27 83  
FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT CARBOHYD 180 180  
FT MOD\_RES 330 330  
FT VARIANT 97 97  
FT VARIANT 239 239  
FT VARIANT 241 241  
FT STRAND 123 126  
FT HELIX 130 134

136 137  
141 148  
158 162  
163 164  
165 166  
175 178  
183 190  
193 197  
198 199  
202 206  
215 219  
227 227  
234 234  
238 240  
241 242  
245 256  
260 266  
267 268  
269 270  
274 276  
280 281  
283 284  
287 296  
297 301  
302 303  
306 312  
313 314  
316 317  
320 324  
330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
SQ SEQUENCE

Query Match 72.9%; Score 1224; DB 1; Length 330;  
Best Local Similarity 97.4%; Pred. No. 7.6e-87;  
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCTCPPEAPEAGAPSVFLFPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNW 139  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 101 KSCDKTHCTCPPEAPELGGPSVFLFPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNW 160

QY 140 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKGKEYCKVSNKALPAPIEKTIS 199  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 161 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKGKEYCKVSNKALPAPIEKTIS 220

QY 200 KAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 259  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 221 KAGQPREPOVYTLPPSREDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 280

QY 260 LDSGDSFFLYSKLTVDKSRWQQGNVFSCSYMHEALHNHYTKQSLSPGK 309  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 281 LDSGDSFFLYSKLTVDKSRWQQGNVFSCSYMHEALHNHYTKQSLSPGK 330

RESULT 2  
GC2\_HUMAN  
ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
AC P01859;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE IG GAMMA-2 CHAIN C REGION.  
GN IGHG2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-82197621; PubMed-6804948;  
RA Ellison J.W., Hood L.E.;  
RT "Linkage and sequence homology of two human immunoglobulin gamma  
heavy chain constant region genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).



[2]  
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IgG2 heavy chain: genetic,  
 RL evolutionary, and functional implications.";  
 RN J. Immunol. 125:1048-1054(1980).  
 [3]  
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
 RX MEDLINE=80001357; PubMed=113060;  
 RA Connell G.E., Parr D.M., Hofmann T.;  
 RT "The amino acid sequences of the three heavy chain constant region  
 RL domains of a human IgG2 myeloma protein.";  
 RN Can. J. Biochem. 57:758-767(1979).  
 [4]  
 RP SEQUENCE OF 238-275 (ZIE).  
 RX MEDLINE=80114419; PubMed=118920;  
 RA Hofmann T., Parr D.M.;  
 RT "A note of the amino acid sequence of residues 381-391 of human  
 RL immunoglobulins gamma chains";  
 RN Mol. Immunol. 16:923-925(1979).  
 [5]  
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
 RX Hofmann T., Parr D.M.;  
 RL Submitted (MAR-1980) to the PIR data bank.  
 [6]  
 RP SEQUENCE OF 1-121 (DOT).  
 RX MEDLINE=9525298; PubMed=7737190;  
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
 RT "Characterization of the two unique human anti-flavin monoclonal  
 RL immunoglobulins.";  
 RN Eur. J. Biochem. 228:886-893(1995).  
 [7]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=72033500; PubMed=4940472;  
 RA Milstein C., Frangione B.;  
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
 RL Biochem. J. 121:217-225(1971).  
 [8]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=69064124; PubMed=5782707;  
 RA Frangione B., Milstein C., Pink J.R.L.;  
 RT "Structural studies of immunoglobulin G.";  
 RL Nature 221:145-148(1969).  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; V00554; CAB58438.1; -.  
 DR PIR; A02148; G2HU.  
 DR MIM; 147110; -.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003600; Ig\_like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igcl; 2.  
 DR SMART; SM00410; Ig-like; 1.  
 DR PROSITE; PS00290; Ig\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT SITE 156 156  
 FT MOD\_RES 326 326  
 FT VARIANT 60 60  
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
 SQ  
 Query Match 69.1%; Score 1159.5; DB 1; Length 326;  
 Best Local Similarity 66.6%; Pred. No. 6.6e-82;  
 Matches 229; Conservative 8; Mismatches 30; Indels 77; Gaps 5;  
 Oy 24 SVAGEQAPGTAPCSRGRSSWSADLDKCMDCASCARPHSDFCGLCAAAP--PAPFRLWRS 81  
 Db 2 STKGPSVFPLAPCSRSTS-----ESTAAIGCLVKDYFFPEPVTWSNS 43  
 Oy 82 -----CDKTH----- 86  
 Db 44 GALTSGVHTFPAVLOSSGLYSLSSVTVVPSNFGTQTYTCNVDRKPSNTKVDKTVVERKCC 103  
 Oy 87 -TCPPCPAPEAGAPSVFLPPKPKDLMISRTPEVTCVVVDVSHEDPEVKENWYDVGVE 145  
 Db 104 VECPPCPAPPVAG-PSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVQFNWYDVGVE 162  
 Oy 146 VHNATKPREOYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIETKISAKGQP 205  
 Db 163 VHNATKPREOYNSTYRVSVLTVLHODWLNKGYCKVSNKGLPAPIETKISAKGQP 222  
 Oy 206 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 265  
 Db 223 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 282  
 Oy 266 PFLYSKLVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPGK 309  
 Db 283 PFLYSKLVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPGK 326  
 RESULT 3  
 GC3\_HUMAN  
 ID GC3\_HUMAN STANDARD; PRT; 290 AA.  
 AC P01860;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).  
 GN IGHG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE (DISEASE PROTEIN WIS).  
 RX MEDLINE=81021548; PubMed=6774747;  
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
 RL gamma 3 heavy-chain disease protein Wis.";  
 RN Biochemistry 19:4304-4308(1980).  
 [2]  
 RP REVISIONS TO 12-97 OF PROTEIN WIS.  
 RX MEDLINE=77118561; PubMed=402363;  
 RA Michaelson T.E., Frangione B., Franklin E.C.;  
 RT "Primary structure of the 'hinge' region of human IgG3. Probable  
 RL quadruplication of a 15-amino acid residue basic unit.";  
 RN J. Biol. Chem. 252:883-889(1977).  
 [3]  
 RP REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).  
 RX MEDLINE=77021516; PubMed=823945;  
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;  
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.  
 RL Structure of the Fc fragment of immunoglobulin G3.";





```

QY 82 -----CDKTH-----TC--PPC 91
DB 46 LTNGVRFPSVRSGLYSLSSVSVTSQPVTCNVAHPATNTKVDKTVAPSTCSKPTC 105
QY 92 PAPAEGAPSVFLPPPKDPTLMTSRPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKT 151
DB 106 PPELLGSPSVFIPPPKDTLMTSRPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKT 165
QY 152 KPREQVNSYRVSVTLVLDHMLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVY 211
DB 166 PLREQQFNSTIRVSVTLVLDHMLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVY 225
QY 212 TLPSREEMTKNOVSLCLVCKFYPSDIWVESNGOPENNYKTPPVLDSDGSGFFLYSK 271
DB 226 TMGPPRELSSRSVSLTCMNGFVPSDISVEWENKGAEDNYKTPPAVLSDGSGFFLYSK 285
QY 272 LTVDKSRWQGNVSCSVMMHEALHNHYTQKSLSLSPGK 309
DB 286 LSVPTSEWQGDVFTCSVMHEALHNHYTQKSLSRSPGK 323

RESULT 6
GC2_CAVPO
ID GC2_CAVPO STANDARD: PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE IG GAMMA-2 CHAIN C REGION.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshrein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). 3. Amino acid sequence of the region around the
half-cysteine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971);
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971);
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
antibodies.";
RL Biochemistry 13:4796-4803(1974);
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;
RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
antibodies.";
RL Biochemistry 13:4804-4811(1974);
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";

```

```

RL Biochemistry 10:26-31(1971).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_GL.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IgC1; 2.
DR SMART: SM00410; Ig_Like; 1.
DR PROSITE: PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 52.2%; Score 876; DB 1; Length 329;
Best Local Similarity 54.1%; Pred. No. 3.4e-60;
Matches 171; Conservative 28; Mismatches 57; Indels 60; Gaps 4;

QY 53 ASCRARPHSDFCLGCAAP--PAPFRLWRS----- 81
DB 14 ASCVDTSGSMNTLGLVKGVFPPEVTVMNSGALTSVHTFPVAVLQSLYSLTSMVTVPS 73
QY 82 -----CDKTH----- 114
DB 74 SQKATCNVAHPASSTKVDKTVETPZBPCTCPKPPENLGGPSPVIFPPKPKDTFLM 133
QY 115 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTPREOYNSTYRVVSVLTVLDH 174
DB 134 ISLTPTVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTPREOYNSTYRVVSVLTVLDH 193
QY 175 WLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGF 234
DB 194 WLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGF 253
QY 235 YPSDIAVESWESNGOP--ENNYKTTTPVLDSDGSEFSLYSLTVDKSRWQGNVSCSVH 292
DB 254 FPADIHWESNRPVSEKEKNTTPPIEDADGSGYFLYSLYSLTVDKSRWQGNVSCSVH 313
QY 293 ALNHYTQKSLSLSPG 308
DB 314 ALNHYTQKSLSLSPG 329

RESULT 7
GC1_MOUSE
ID GC1_MOUSE STANDARD: PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).

```

|   |  |  |                                  |
|---|--|--|----------------------------------|
| Best Local Similarity 50.0%; Pred. No. 2.5e-57;             |  | Matches 163; Conservative 55; Mismatches 79; Indels 29; Gaps |                                  |
| QY  | 1 MARGSLRRLRLVLGLWLALLRVSAGEQAPGTAPCSRGSWS-----ADLDKCM 50            |  |                                  |
| Db  | 11 LAPGSAQTNSMTLG---CLVKGYFPE--PVTYVNSGSLSSGVHTFPVAVLQSDLYTLS 65     |  |                                  |
| QY  | 51 DCASCRARPHSDFCLGCAAAAPPAPF-----RLLRSCDKTHTCTPP--CPAPEAEGAPSVF 100 |  |                                  |
| Db  | 66 SSVTPSPSPRSETVTNCVAHPASSTVKDKKIVPRDCG----CKPCICTVPEVS---SVF 111   |  |                                  |
| QY  | 104 LFPPKPKDTLMISTPEVTCVVVDVSDHEDPEVKFNWYGVGVEHNATKPREEOYNSTYR 160   |  |                                  |
| Db  | 119 IFPPKPKDVLITITLTPKVTCTVVDISKDDPEQSFVSWFDDVEVHTAQTQPREQFNSTER 177 |  |                                  |
| QY  | 164 VVSVLTVHLDWLNKYEKKVCKVSNKALPAPTEKTSKAGQPRPQVYVTLPPSPREEMTKN 220  |  |                                  |
| Db  | 179 SVSELPIMHLDWLNKYEKKVCKVSNKALPAPTEKTSKAGQPRPQVYVTLPPSPREEMTKN 236 |  |                                  |
| QY  | 224 QVSLTCLVKGYFSDIAVENESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKRSWQOQN 280  |  |                                  |
| Db  | 239 KVSLTCTMITDFEFDIIEVQWQNPQAEYKNTQPIIMNTNGSYEYVSKLVQKSNWEAGN 299   |  |                                  |
| QY  | 284 VFCSVMHEALHNHYTOKSLSLSPGK 309                                    |  |                                  |
| Db  | 299 TFCVSLHEGLHNHHTKSLSHSPGK 324                                     |  |                                  |
| RESULT 8  |  |  |                                  |
| ICL_RATE  | ICL_RATE   | STANDARD;  | PRT; 326 AA.                     |
| AC  | P20759;  |  |                                  |
| DT  | 01-FEB-1991 (Rel. 17, Created)                                       |  |                                  |
| DT  | 01-FEB-1991 (Rel. 17, Last sequence update)                          |  |                                  |
| DE  | 15-JUL-1999 (Rel. 38, Last annotation update)                        |  |                                  |
| DE  | IG GAMMA-1 CHAIN C REGION.   |  |                                  |
| OS  | Rattus norvegicus (Rat).   |  |                                  |
| OC  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |  |                                  |
| OC  | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus |  |                                  |
| ON  | NCBI_TaxID=10116;  |  |                                  |
| RX  | [1]  |  |                                  |
| RP  | SEQUENCE FROM N.A.   |  |                                  |
| RP  | MEDLINE=89232738; PubMed=3149946;                                    |  |                                  |
| RA  | Brueggemann M.   |  |                                  |
| RA  | "Evolution of the rat immunoglobulin gamma heavy-chain gene family." |  |                                  |
| RL  | Gene 74:473-482(1988).   |  |                                  |
| DR  | PIR; PS0017; PS0017.   |  |                                  |
| DR  | InterPro; IPR003006; Ig_MHC.   |  |                                  |
| DR  | InterPro; IPR003597; Ig_cl.  |  |                                  |
| DR  | Pfam; PF00047; Ig; 3.  |  |                                  |
| DR  | SMART; SM00407; Igcl; 2.   |  |                                  |
| DR  | PROSITE; PS00290; IG_MHC; 1.   |  |                                  |
| KW  | Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.        |  |                                  |
| FT  | NON_TER 1 97   |  | CHI.                             |
| FT  | DOMAIN 1 97  |  | HINGE.                           |
| FT  | DOMAIN 98 112  |  | CH2.                             |
| FT  | DOMAIN 113 219   |  | CH3.                             |
| FT  | DOMAIN 220 326   |  |                                  |
| FT  | DISULFID 27 82   |  | INTERCHAIN (WITH A HEAVY CHAIN). |
| FT  | DISULFID 102 102   |  | INTERCHAIN (WITH A HEAVY CHAIN). |
| FT  | DISULFID 106 106   |  | INTERCHAIN (WITH A HEAVY CHAIN). |
| FT  | DISULFID 109 109   |  | INTERCHAIN (WITH A HEAVY CHAIN). |
| FT  | DISULFID 111 111   |  | INTERCHAIN (WITH A HEAVY CHAIN). |
| FT  | DISULFID 140 200   |  |                                  |
| FT  | DISULFID 246 304   |  |                                  |
| FT  | CARBOHYD 176 176   |  |                                  |
| QY  | SEQUENCE 326 AA; 35946 MW; 0139AB45EF49B9DA CRC64;                   |  |                                  |
| Query Match 49.9%; Score 838; DB 1; Length 326;             |  |  |                                  |
| Best Local Similarity 60.5%; Pred. No. 2.8e-57;             |  |  |                                  |
| Matches 150; Conservative 43; Mismatches 47; Indels 8; Gaps |  |  |                                  |

Qy 67 CAAAPPAPF-----RLLRSCDKTHTCCPAPAEAGAPSVFLPPPKDLMISRTPEV 121  
Db 82 CNVAHPASSTVKKIYPRNGC--GCKPCICTGSE--VSSVFIPPPKPKDVLTLTPKV 138  
Qy 122 TCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSLVTLVHODWLNKGEY 181  
Db 139 TCVVVDISQDDPEVHFVSWFVNDKEVHTAQTTPPEEQFNSTERSVSELPILHODWLNKGRTF 198  
Qy 182 CKKYSNKALPAPIKTKISKAGOPREPOVYTLPPSRREMTKNQVSLCLVKGFPSPDIAY 241  
Db 199 RCKYTSAAFPSPIEKTKISKEPGRTQVPHVYTPSTPKEMTQNEYSITCMWKGFPDPDIY 258  
Qy 242 EWESNGPNENKYPPTPVLDSDGFFLYSKLTVDKSRWQGNFVSCSVMHHEALNHHYTK 301  
Db 259 EWQNGQPOENKYPPTMTDGDGFFLYSKLVNKKVKEKQOQNTFTCSVLHHEGLNHHTEK 318  
Qy 302 SLSLSPGK 309  
Db 319 SLSHSPGK 326

## RESULT 9

GC3\_MOUSE ID GC3\_MOUSE STANDARD: PRT: 329 AA.  
AC P22436;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85027161; PubMed=6092053;  
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
RA Tucker P.W., Blattner F.R.;  
RT "Structural analysis of the murine IgG3 constant region gene.";  
RL EMBO J. 3:2041-2046(1984).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; J00451; -, NOT\_ANNOTATED\_CDS.  
DR PIR; B02156; G3WSC.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003597; Ig\_cl.  
DR InterPro; IPR003600; Ig-like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGcl; 2.  
DR SMART; SM00410; IG-like; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Transmembrane; Alternative splicing.  
FT NON\_TER 1 1  
FT DOMAIN 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 223 CH2.  
FT DOMAIN 224 327 CH3.  
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 49.8%; Score 836; DB 1; Length 329;  
Best Local Similarity 62.7%; Pred. No. 4e-57;  
Matches 156; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

Qy 67 CAAAPPAPFRLLRSCD-----KTHTCP--PCPAEAGAPSVFLPPPKDLMISRTPE 120

Db 81 CNVAHPASSTKELIKRIEPIKPSPTPGSCPPGNILGSPSVFIPPPKPKDMLSLTPK 140  
Qy 121 VTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSLVTLVHODWLNKGE 180  
Db 141 VTCVVVDVSEDDPDVHVSFWFVNDKEVHTAQTTPPEEQFNSTERSVSELPILHODWLNKGRKE 200  
Qy 181 YKCVSNKALPAPIKTKISKAGOPREPOVYTLPPSRREMTKNQVSLCLVKGFPSPDIAY 240  
Db 201 FKCKVNNKALPAPIERTLSKPKRAQTQVYTIPTPREQMSKKVSLTCLVTFNSEALS 260  
Qy 241 VEWESNGPNENKYPPTPVLDSDGFFLYSKLTVDKSRWQGNFVSCSVMHHEALNHHYTK 300  
Db 261 VEWERNGELEQDYKNTPTPILSDGTYFLYSLKLTVDTSWLOGEITCSVWHEALNHHHTQ 320  
Qy 301 KSLSLSPGK 309  
Db 321 KNLRSRSPGK 329

## RESULT 10

GC1M\_MOUSE ID GC1M\_MOUSE STANDARD: PRT: 393 AA.  
AC P01869;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=80045036; PubMed=115593;  
RA Honjo T., Ohta M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
RA Takahashi N., Mano Y.;  
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
RT gamma 1 chain gene";  
RL Cell 18:559-568(1979).  
RN [2]  
RP SEQUENCE OF 323-393 FROM N.A.  
RX MEDLINE=82197626; PubMed=6804950;  
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;  
RT "mRNA for surface immunoglobulin gamma chains encodes a highly  
RT conserved transmembrane sequence and a 28-residue intracellular  
RT domain";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).  
RN [3]  
RP SEQUENCE OF 323-366 FROM N.A.  
RX MEDLINE=82115295; PubMed=6799207;  
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,  
RA Eisenberg D., Wall R.;  
RT "Gene segments encoding transmembrane carboxyl terminl of  
RT immunoglobulin gamma chains";  
RL Cell 26:19-27(1981).  
RN [4]  
RP SEQUENCE OF 1-44 FROM N.A.  
RX MEDLINE=822212190; PubMed=6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
RT immunoglobulin gamma chains";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
DR EMBL; V00793; CAA24172.1; -.
DR EMBL; V00793; CAA24173.1; -.
DR EMBL; V00793; CAA24174.1; -.
DR PIR; B02159; G1MSM.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 INTERCHAIN (WITH A HEAVY CHAIN).
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;
```

```
Query Match 49.7%; Score 833.5; DB 1; Length 393;
Best Local Similarity 49.8%; Pred. No. 7.6e-57;
Matches 162; Conservative 55; Mismatches 79; Indels 29; Gaps 7;

QY 1 MARGSLRLLRLVLLGLWALRLSVAGEAQTAPCSRGSWS-----ADLDKCM 50
DB 11 LAPGSAATNSMTLIG---CLVKGYPE--PVTVTWNSGLSSGVHTFFAVLQSLDTLS 65
QY 51 DCASCRARPHSDFCLGCAAPPAPE-----RLLRSCDKTHTCPP--CPAPAEAGPSVF 103
DB 66 SSVTPSSRPSETVTCNVAPASSTKVKIIPRDCG---CKPICVPEVS---SVF 118
QY 104 LFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYR 163
DB 119 IFFPKPKDVLITLTPKVTCCVVVDISKDDPEVQFSWFVDVVEHTAQTQPREEQFNSTR 178
QY 164 VVSVLTVLHODWLNKKEYCKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKN 223
DB 179 SVSELPIMHODWLNKKEFKCRYNRAAFAPAIETKISKTRKAPQVYTIPTPKQEQAKD 238
QY 224 QVSLTCLVKGFVPSDIAVEWESNGQPENNYKTPPPVLDSDGSFELYKLTVDKSRWQGN 283
DB 239 KVSLTCTMTDFPEDITVSWQNGQPAENYKNTQPIIMNTNGSFYVYSKLNVKRSNWEAGN 298
QY 284 VFSCSYMHEALHNHYTQKSLSLSPG 308
DB 299 TITCSVLHESLHNHHEKSLSHSPG 323

RESULT 11
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
FT SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;
```

```
Query Match 49.2%; Score 825; DB 1; Length 333;
Best Local Similarity 64.0%; Pred. No. 2.8e-56;
Matches 146; Conservative 33; Mismatches 49; Indels 0; Gaps 0;

QY 82 CDKTHTCPPCPAPEAGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 141
DB 106 CPTCPTCHKCPVPELLGGPSVFIFFPKPKDILLISQNAKVTCTVVDVSEEDPVQFSWF 165
QY 142 DGVVEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKA 201
DB 166 NNVEVTAQTQPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKP 225
QY 202 KQPREPVQVYTLPPSREEMTKNQVSLTCLVKGFVPSDIAVEWESNGQPENNYKTPPVLD 261
DB 226 KGLVRPQVYVMGPPTTEQLTEQVSLTCLSGFLPNDIGVETWSNGHIEKNYKNTPEVMD 285
QY 262 SDGSFELYSLKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 309
DB 286 SDGSFEMYSKLVNRSRWDSDRAPFCVSVHGLHNNHVEKSISSRPCK 333
```

```
RESULT 12
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987.
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
```

RA Komatomy M., Clayton L., Rogers J., Robertson S., Kettman J.,  
RA Wall R.;  
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene  
RT segment.";  
RL Nucleic Acids Res. 11:6775-6785(1983).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J00451; AAB59655.1; -;  
DR EMBL: V01526; CAA24767.1; ALT\_SEQ.  
DR PIR: A02155; G3MSM.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003597; Ig\_cl.  
DR InterPro: IPR003600; Ig\_like.  
DR Pfam: PF00047; Ig; 3.  
DR SMART: SM00407; Igcl; 2.  
DR SMART: SM00410; Ig\_like; 1.  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Transmembrane; Alternative splicing.  
FT NON\_TER 1 1  
FT DOMAIN 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 223 CH2.  
FT DOMAIN 224 327 CH3.  
FT TRANSMEM 346 362 POTENTIAL.  
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
FT CONFLICT 333 333 E -> G (IN REF. 2).  
FT CONFLICT 342 342 E -> Q (IN REF. 2).  
FT CONFLICT 388 388 P -> F (IN REF. 2).  
SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 49.2%; Score 825; DB 1; Length 398;  
Best Local Similarity 62.3%; Pred. No. 3.4e-56;  
Matches 154; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

QY 67 CAAPAPAPFLLWRSCLD-----KTHTCP--PCPAEAGAPSVFLPPKPKDTLMISRTPE 120  
DB 81 CNVAHPASKTELKRIEPRPKPSTPPGSSCPPGNIILGSPVFIFFPKPKDALMISLTPK 140  
QY 121 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKE 180  
DB 141 VTCVVVDVSEDDPDVHVSWFVFNKVEHTAWTQPREAQYNSTFRVVSALPILQHQQDMRGKE 200  
QY 181 YKCVSNKALPAPIETKISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIA 240  
DB 201 FKCKVNNKALPAPERTISKRAQTPQVYTIIPPREQMSKKKVSITCLVITNFFSEASIS 260  
QY 241 VEWESNGPENNYKTPPVLDSGSEFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQ 300  
DB 261 VEWERNELQDYKNTPIPLDSGTGYFLYKSLTVDTSLWLGQEIFTCVSVHEALHNHTQ 320  
QY 301 KSLSLSP 307  
DB 321 KNLRSRP 327

RESULT 13  
GCC\_RAT ID GCC\_RAT STANDARD; PRT; 329 AA.  
AC P20762;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2C CHAIN C REGION.  
OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88166903; PubMed=3127222;  
RA Brueggemann M., Delmasro-Galfre P., Waldmann H., Calabi F.;  
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
RT region cDNA: extensive homology to mouse gamma 3.";  
RL Eur. J. Immunol. 18:317-319(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X07189; CAA30169.1; -;  
DR PIR: S00847; S00847.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003597; Ig\_cl.  
DR InterPro: IPR003600; Ig\_like.  
DR Pfam: PF00047; Ig; 3.  
DR SMART: SM00407; Igcl; 2.  
DR SMART: SM00410; Ig\_like; 1.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1 1  
FT DOMAIN 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 222 CH2.  
FT DOMAIN 223 329 CH3.  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 143 203  
FT DISULFID 249 307  
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 48.7%; Score 817; DB 1; Length 329;  
Best Local Similarity 53.8%; Pred. No. 1.1e-55;  
Matches 155; Conservative 46; Mismatches 49; Indels 38; Gaps 3;

QY 33 TAPCSRSSWS-----ADLDKMDCCASCARPHSDFCLGCAAPAPFRLWRS 81  
DB 69 TVPSSTWSSQTVTCVAHPATKSNLIKRIEPRPKRPPTDIC-----S 112  
QY 82 CDKTHTCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 141  
DB 113 CD-----DNLGRPSVFIFFPKPKDTLITLPKVTVCVVVDVSEEDPQVSWFV 161  
QY 142 DGEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISK 201  
DB 162 DNVKFTAQTPQHEEQUNGITFRVYSTLHIHQDWMGSEKFKCKVNNKDLPSFIETISK 221  
QY 202 KGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGPENNYKTPPPVLD 261  
DB 222 RGKARTPQVYTIIPPREQMSKKKVSITCLVITNFFSEASISVEMERNGELEDYKNTLPVLD 281  
QY 262 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309  
DB 282 SDESIFLYSKLSDVTDSDWMRGDIYTCVSVHEALHNHTQKSLSLSPGK 329

RESULT 14  
GCC\_RAT ID GCC\_RAT STANDARD; PRT; 322 AA.  
AC P20760;  
DT 01-FEB-1991 (Rel. 17, Created)



```
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 47.9%; Score 804.5; DB 1; Length 322;
Best Local Similarity 47.0%; Pred. No. 1e-54;
Matches 157; Conservative 38; Mismatches 62; Indels 77; Gaps 5;

Qy 30 APGTAPCRGSSWADLCKMDCASCRAHPHSDFCLGCAAP--PAPFRLWRS----- 81
||||| :| ||| ||| :| |
Dy 12 APGTA-----LKNSMVTGLCLVKGYFPEPVTVTVNSGALSSG 49
||||| :| ||| ||| :| |
Qy 82 -----CDKTH-----TCPPCPAPE 95
||||| :| ||| ||| :| |
Dy 50 VHTFPVAVLQSLYLTSTVTPSWSSQAVTCNVHPASSTKYVDKKIVPRECNPCGCTG 109
||||| :| ||| ||| :| |
Qy 96 AEGAPSVLFPPKPKDPTLMIISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKATKPRE 155
:| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 110 SE-VSSVFIFPPKTKDVLITLTKPVTCVVVDVSDQNDPEVFESWFDIDVEVHTAQTAAPE 168
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 156 EYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQVYTTLP 215
:| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 169 KOSNTLSRVSSELPVHVRDLWNGKTKCKVNSGAPPAPPIEKISKEGTPRGPQVYTMAP 228
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 216 SREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPTPVLDSDGSFYSKLTVD 275
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 229 PREMTQSQVITCKVKGYPDPDIYETWMKNGQPOENTKNTPTTMDTGSGYFLSKLVNK 288
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 276 KSRWQQGNVFCVSMVEALHNYHTQKLSLSPGK 309
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 289 KETWQQGNTFTCSVLHLEGLHNHTTKLSLSPGK 322
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|

RESULT 15
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE
```

```
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN Nucleic Acids Res. 8:3143-3155(1980).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
RT and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RN Nucleic Acids Res. 9:1365-1381(1981).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RT family.";
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN (4)
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin:amino-acid sequence of the Fc fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
RN (5)
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RL immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -
DR PIR; A02152; G2MSA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
```

```
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 47.6%; Score 799.5; DB 1; Length 330;
Best Local Similarity 49.7%; Pred. No. 2.5e-54;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;

QY 55 CRARPHSDFCGLCAAP--PAPFRLLR----- 81
Db 15 CGDTTGSSTLGLVKGYPFVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSST 74
QY 82 -----CDKTH-----TCPP--CPAPEAGAPSVLFPPKPKDVL 113
Db 75 WPSQSTICNVAHASSTKVDKIEPRGPTIKPCPPCKPAPNLLGGPSVFIFFPKKDVL 134
QY 114 MISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVWSVLTVLHQ 173
Db 135 MISLPIVTCVVVDVSEDDPDVQISWFEVNNVEVHTAQTQTHREDYNSTLRVVSALPIQHQ 194
QY 174 DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYITLPPSRREEMTKNQVSLTCLVKG 233
Db 195 DMSGKEFKCKVNNKDLAPDIERTISKPGSVRAPQVYVLPPEPEEMTKKQVTLTCMVD 254
QY 234 FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSVNHEA 293
Db 255 FMPEDIVETWNNKGTLELNAKTEPVLDSDGSYFMYSKLRVEKKNWVERNSYSCSVVHEG 314
QY 294 LHNHYTOKSLSLSPGK 309
Db 315 LNHHTTKFSFRTPGK 330
```

Search completed: March 11, 2002, 15:23:50  
Job time: 574 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:12:36 ; Search time 31.92 seconds  
(without alignments)  
737.404 Million cell updates/sec

Title: US-09-742-454A-7  
Perfect score: 1678  
Sequence: 1 MARGSLRRLRLVGLWLA.....MHEALHNHYTKQSLSLSPGK 309

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID       | Description         |
|------------|--------|-------------|--------|----------|---------------------|
| 1          | 1230   | 73.3        | 374    | 2 S69339 | Ig heavy chain V r  |
| 2          | 1224   | 72.9        | 255    | 4 S31866 | Ig gamma-1 chain C  |
| 3          | 1224   | 72.9        | 330    | 1 GHU    | Ig gamma-1 chain C  |
| 4          | 1171.5 | 69.8        | 377    | 2 A23511 | Ig gamma-3 chain C  |
| 5          | 1171   | 69.8        | 234    | 2 P70207 | Ig gamma chain C r  |
| 6          | 1169.5 | 69.7        | 377    | 2 A60764 | Ig gamma-3 chain C  |
| 7          | 1159.5 | 69.1        | 326    | 1 G2HU   | Ig gamma-2 chain C  |
| 8          | 1146.5 | 68.3        | 289    | 1 G3HWT  | Ig gamma-3 heavy c  |
| 9          | 1139   | 67.9        | 327    | 1 G4HU   | Ig gamma-4 chain C  |
| 10         | 928    | 55.3        | 328    | 2 I47159 | Ig gamma 2a chain C |
| 11         | 923    | 55.0        | 328    | 2 I47160 | Ig gamma 2b chain   |
| 12         | 910.5  | 54.3        | 323    | 1 GHRB   | Ig gamma chain C r  |
| 13         | 907    | 54.1        | 277    | 2 I47162 | Ig gamma 4 chain c  |
| 14         | 895    | 53.3        | 328    | 2 I47158 | Ig gamma 1 chain c  |
| 15         | 893    | 53.2        | 328    | 2 I47161 | Ig gamma 3 chain c  |
| 16         | 876    | 52.2        | 329    | 1 G2GP   | Ig gamma-2 chain C  |
| 17         | 866    | 51.6        | 470    | 2 S22080 | Ig heavy chain pre  |
| 18         | 842    | 50.2        | 444    | 2 PC4436 | monoclonal antibody |
| 19         | 842    | 50.2        | 472    | 2 S31459 | Ig gamma-1 chain -  |
| 20         | 840    | 50.1        | 308    | 2 C30554 | Ig heavy chain C r  |
| 21         | 838.5  | 50.0        | 324    | 1 GLMS   | Ig gamma-1 chain C  |
| 22         | 838    | 49.9        | 326    | 2 PS0017 | Ig gamma-1 chain C  |
| 23         | 836    | 49.8        | 329    | 1 G3MSC  | Ig gamma-3 chain C  |
| 24         | 833.5  | 49.7        | 393    | 1 G1MSM  | Ig gamma-1 chain C  |
| 25         | 825    | 49.2        | 333    | 2 PS0018 | Ig gamma-2b chain   |
| 26         | 825    | 49.2        | 398    | 1 G3MSM  | Ig gamma-3 chain C  |
| 27         | 817    | 48.7        | 329    | 2 S00847 | Ig gamma-2c chain   |
| 28         | 804.5  | 47.9        | 322    | 2 PS0019 | Ig gamma-2a chain   |
| 29         | 799.5  | 47.6        | 330    | 1 G2MSA  | Ig gamma-2a chain   |

ALIGNMENTS

RESULT 1

S69339  
Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
C:Accession: S69339; S72664  
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; MUID:95262687  
A:Accession: S69339  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <KHA>  
A:Cross-references: EMBL:X81695  
R:Khamilichi, A.A.  
submitted to the EMBL Data Library, September 1994  
A:Reference number: S72664  
A:Accession: S72664  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-140,'C',142-374 <KH2>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

|                       |        |   |               |                      |
|-----------------------|--------|---|---------------|----------------------|
| Query Match           | 73.3%  | Score 1230;   | DB 2;         | Length 374;          |
| Best Local Similarity | 98.3%; | Pred. No. 2.3e-81;  |               |                      |
| Matches               | 226;   | Conservative 1;   | Mismatches 3; | Indels 0; Gaps 0;    |
| QY                    | 80     | RSCDTHTCPPCPAPEAGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW | 139           | Ig gamma-2a chain    |
| Db                    | 145    | KSCDTHTCPPCPAPEAGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW | 204           | Ig gamma-2a chain    |
| QY                    | 140    | YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHDLWLNKCKVSKNKPAPIEKTTIS    | 199           | Ig gamma-2b chain    |
| Db                    | 205    | YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHDLWLNKCKVSKNKPAPIEKTTIS    | 264           | Ig gamma-2b chain    |
| QY                    | 200    | KAKGPREQVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV  | 259           | Ig gamma-1 chain V-I |
| Db                    | 265    | KAKGPREQVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV  | 324           | Ig mu chain C regi   |
| QY                    | 260    | LSDSGSFFLYSKLTVDKSRWQQGNVFCSPVMEALHNHYTKQSLSLSPGK         | 309           | Ig epsilon chain C   |
| Db                    | 325    | LSDSGSFFLYSKLTVDKSRWQQGNVFCSPVMEALHNHYTKQSLSLSPGK         | 374           | Ig mu heavy chain    |

RESULT 2

S31866  
Ig gamma-1 chain C region - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C;Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C;Accession: S31866

R;Filpula, D.  
submitted to the EMBL Data Library, February 1993  
C;Description: Screening method for protein-protein interactions of cloned gene products.  
A;Reference number: S31866  
A;Molecule type: mRNA  
A;Accession: S31866  
A;Residues: 1-255 <FIL>  
A;Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069  
C;Keywords: immunoglobulin  
F;1-22/Region: Escherichia coli outer membrane protein A precursor  
F;23-255/Region: human Ig gamma-1 chain C region

Query Match 72.9%; Score 1224; DB 4; Length 255;  
Best Local Similarity 97.4%; Pred. No. 4e-81;  
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKHTCCPCPAEAGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
Db :|||||  
26 KSCDKHTCCPCPAELLGSPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 85  
QY 140 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
Db :|||||  
86 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 145  
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259  
Db :|||||  
146 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 205  
QY 260 LDSGSEFLYSLKTVDSKRWQGVNFVCSVMHEALHNYTKQSLSLSPGK 309  
Db :|||||  
206 LDSGSEFLYSLKTVDSKRWQGVNFVCSVMHEALHNYTKQSLSLSPGK 255

## RESULT 3

Ig gamma-1 chain C region - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999  
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146  
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A;Reference number: A93433; MUID:82274238  
A;Accession: A93433  
A;Molecule type: DNA  
A;Residues: 1-330 <ELL>  
A;Cross-references: EMBL:Z17370  
A;Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) markers.  
A;Note: Lys-330 is removed after translation  
R;Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A;Reference number: S33904  
A;Accession: S36861  
A;Molecule type: DNA  
A;Residues: 2-330 <HAR>  
A;Cross-references: EMBL:Z17370  
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a  
A;Reference number: S33887; MUID:83001943  
A;Accession: S33887  
A;Molecule type: DNA  
A;Residues: 88-113;235-330 <TAK>  
A;Cross-references: EMBL:Z17370  
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen  
A;Reference number: A90563; MUID:71064024  
A;Contents: myeloma protein Eu  
A;Accession: B90563

A;Molecule type: protein  
A;Residues: 1-96,'R',98-135 <CUN>  
A;Note: this sequence has the Gln(3) marker, 97-Arg  
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se  
A;Reference number: A90564; MUID:71064025  
A;Contents: Eu  
A;Accession: A90564  
A;Molecule type: protein  
A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2  
A;Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met  
R;Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni  
igen Primaerstruktur.  
A;Reference number: A91668; MUID:77070269  
A;Contents: myeloma protein Nie  
A;Accession: B91668  
A;Molecule type: protein  
A;Residues: 134,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E'  
A;Note: this sequence has the Gln(17) and Gln(1) markers  
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1  
A;Reference number: A91723; MUID:83289131  
A;Contents: myeloma protein KOL; disulfide bonds  
A;Accession: A91723  
A;Molecule type: protein  
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <  
A;Note: this sequence has the Gln(3) and Gln(non-1) markers  
R;Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disul  
A;Reference number: A90565; MUID:71064027  
A;Contents: annotation; disulfide bonds  
R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunog  
enamide cleavage products, and the disulfide bridges.  
A;Reference number: A91667; MUID:77070267  
A;Contents: annotation; disulfide bonds  
C;Genetics:  
A;Gene: GDB:IGHG1  
A;Cross-references: GDB:120085; OMIM:147100  
A;Map position: 14q32.33-14q32.33  
A;Introns: 99/1; 114/1; 224/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( ha  
in disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;20-85/Domain: immunoglobulin homology <IM1>  
F;137-206/Domain: immunoglobulin homology <IM2>  
F;243-310/Domain: immunoglobulin homology <IM3>  
F;27-83,144-204,250-308/Disulfide bonds: #status experimental  
F;103/Disulfide bonds: interchain (to light chain) #status experimental  
F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 72.9%; Score 1224; DB 1; Length 330;  
Best Local Similarity 97.4%; Pred. No. 5.4e-81;  
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKHTCCPCPAEAGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 139  
Db :|||||  
101 KSCDKHTCCPCPAELLGSPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 160  
QY 140 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199  
Db :|||||  
161 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 220  
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259

|||||  
Db 221 KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 280  
Qy 260 LDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 309  
Db 281 LDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 330

RESULT 4  
A23511  
Ig gamma-3 chain C region (allotype G3m(b)) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
C:Accession: A23511  
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
Nucleic Acids Res. 14, 1779-1789, 1986  
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: ccd  
A:Reference number: A23511; MUID:86148507  
A:Accession: A23511  
A:Molecule type: DNA  
A:Residues: 1-377 <HUC>  
A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 69.8%; Score 1171.5; DB 2; Length 377;  
Best Local Similarity 83.0%; Pred. No. 3.8e-77;  
Matches 220; Conservative 9; Mismatches 33; Indels 3; Gaps 1;  
Qy 48 KCMDCASCARPHSDFC---LGCAAPAPFLLWRSCKDTHCPCPAPAEAGAPSVFL 104  
Db 113 RCPEPKSCDTPPPCPCEPKSCDTPPPCPCEPKSCDTPPPCPAPAEAGAPSVFL 172  
Qy 105 PPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREEQYNSTYRV 164  
Db 173 PPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREEQYNSTYRV 232  
Qy 165 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSREEMTKNQ 224  
Db 233 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSREEMTKNQ 292  
Qy 225 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSGSEFLYSLKLTVDKSRWQOGNV 284  
Db 293 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSGSEFLYSLKLTVDKSRWQOGNV 352

RESULT 5  
PT0207  
Ig gamma chain C region - chimpanzee  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
C:Accession: PT0207  
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.  
Mol. Immunol. 28, 319-322, 1991  
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
A:Reference number: PT0207; MUID:91287716  
A:Accession: PT0207  
A:Molecule type: mRNA  
A:Residues: 1-234 <EHR>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 69.8%; Score 1171; DB 2; Length 234;  
Best Local Similarity 96.0%; Pred. No. 2.4e-77;  
Matches 214; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 80 RSCDKTHTCPCPAPAEAGAPSVFLPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNW 139  
Db 12 KSCDTHTCPCCAPELLGSPSVFLPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNW 71  
Qy 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 199  
Db 72 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 131  
Qy 200 KAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 259  
Db 132 KAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 191  
Qy 260 LDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKS 302  
Db 192 LDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKS 234

RESULT 6  
A60764  
Ig gamma-3 chain C region, form LAT - human  
C:Species: Homo sapiens (man)  
C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999  
C:Accession: A60764  
R:Huck, S.; Lefranc, G.; Lefranc, M.P.  
Immunogenetics 30, 250-257, 1989  
A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conv  
A:Reference number: A60764; MUID:90007613  
A:Accession: A60764  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-377 <HUC>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 69.7%; Score 1169.5; DB 2; Length 377;  
Best Local Similarity 83.0%; Pred. No. 5.3e-77;  
Matches 220; Conservative 9; Mismatches 33; Indels 3; Gaps 1;  
Qy 48 KCMDCASCARPHSDFC---LGCAAPAPFLLWRSCKDTHCPCPAPAEAGAPSVFL 104  
Db 113 RCPEPKSCDTPPPCPCEPKSCDTPPPCPCEPKSCDTPPPCPAPAEAGAPSVFL 172  
Qy 105 PPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREEQYNSTYRV 164  
Db 173 PPPPKDGLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREEQYNSTYRV 232  
Qy 165 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSREEMTKNQ 224  
Db 233 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSREEMTKNQ 292  
Qy 225 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSGSEFLYSLKLTVDKSRWQOGNV 284  
Db 293 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSGSEFLYSLKLTVDKSRWQOGNV 352  
Qy 285 FSCVMHEALHNHYTQKSLSLSPGK 309  
Db 353 FSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 7  
G2H0  
Ig gamma-2 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
C:Accession: A93906; A92809; A90752; A93132; A02148

R;Ellison, J.; Hood, L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
A:Reference number: A93906; MUID:82197621  
A:Accession: A93906  
A:Molecule type: DNA  
A:Residues: 1-326 <ELL>  
A:Cross-references: GB:V00554; GB:J00230; NID:932759; PIDN:CAB58438.1; PID:g6066056  
A:Note: Lys-326 is probably removed posttranslationally  
R;Wang, A.C.; Tung, E.; Eudenberg, H.H.  
J. Immunol. 125, 1048-1054, 1980  
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f  
A:Reference number: A92809; MUID:81007873  
A:Contents: myeloma protein T11  
A:Accession: A92809  
A:Molecule type: protein  
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
A:Note: Trp-156 is at or near the complement-binding site  
R;Connell, G.E.; Parr, D.M.; Hofmann, T.  
Can. J. Biochem. 57, 758-767, 1979  
A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
A:Reference number: A90752; MUID:80001357  
A:Contents: myeloma protein Z1e  
A:Accession: A90752  
A:Molecule type: protein  
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, '222', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
A:Note: this sequence has since been revised  
R;Hofmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
A:Reference number: A93132; MUID:80114419  
A:Contents: Z1e  
A:Accession: A93132  
A:Molecule type: protein  
A:Residues: 238-275 <HOF>  
R;Hofmann, T.; Parr, D.M.  
Submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation: Z1e, revisions to residues 25, 59, 60, and 264-268  
A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
ned  
R;Millstein, C.; Frangione, B.  
Biochem. J. 121, 217-225, 1971  
A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.  
A:Reference number: A90253; MUID:72033500  
A:Contents: annotation: myeloma protein Sa, disulfide bonds  
R;Frangione, B.; Milstein, C.; Pink, J.R.L.  
Nature 221, 145-148, 1969  
A:Title: Structural studies of immunoglobulin G.  
A:Reference number: A93157; MUID:69064124  
A:Contents: annotation; Sa, disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGHG2  
A:Cross-references: GDB:119338; OMIM:147110  
A:Map position: 14q32.33-14q32.33  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;20-85/Domain: immunoglobulin homology <IMI>  
F;133-202/Domain: immunoglobulin homology <IM2>  
F;239-306/Domain: immunoglobulin homology <IM3>  
F;14/Disulfide bonds: Interchain (to light chain) #status experimental  
F;27-83,140-200,246-304/Disulfide bonds: #status experimental  
F;102,103,106,109/Disulfide bonds: Interchain (to heavy chain) #status experimental  
F;176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.1%; Score 1159.5; DB 1: Length 326;  
Best Local Similarity 66.8%; Pred. No. 2.3e-76;  
Matches 229; Conservative 8; Mismatches 30; Indels 77; Gaps 5;  
QY 24 SVAGEAPGTPAGCSRGSSWSADLDKMCDCASCRPHSFCLGCAAP--PAPFRLIWR 81

Db 2 STKGPSVFPLAPCSRSTS-----ESTAAAGCLVKDYFPEPVTVSWNS 43  
QY 82 -----CDKTH----- 86  
Db 44 GALTSGVHTPEAVLQSSGLYSLSVTVTPSSNFGTQTYTCNVDPHKPSNTKVDKTVKRC 103  
QY 87 -TCPGPCPAEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVE 145  
Db 104 VECPCPAPPVAG-PSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNKYVDGVE 162  
QY 146 VHNAKTKPREEQNSTYRVVSVLTVLHQLDNLNGKEYKCKVSNKALPAPIETKTISKAKGP 205  
Db 163 VHNAKTKPREEQNSTFRVSVLTVVHQLDNLNGKEYKCKVSNKGLPAPIETKTISKTKGP 222  
QY 206 REPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 265  
Db 223 REPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 282  
QY 266 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309  
Db 283 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326  
RESULT 8  
G3HUWI  
Ig gamma-3 heavy chain disease proteins - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
C:Accession: A90442; A92219; A90198; A93915; A02149  
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
Biochemistry 19, 4304-4308, 1980  
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 hea  
A:Reference number: A90442; MUID:81021548  
A:Contents: heavy chain disease protein Wis  
A:Accession: A90442  
A:Molecule type: protein  
A:Residues: 1-289 <FRA>  
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interch  
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12  
A:Note: the sequence of residues 42-76 was taken from the reference that follows  
R;Michaelsen, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplicat  
A:Reference number: A92219; MUID:77118561  
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protei  
A:Accession: A92219  
A:Molecule type: protein  
A:Residues: 12-97 <MIC>  
A:Note: the hinge region in gamma-3 chains is about four times as long as in other ga  
idue segment (12-28)  
A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form in  
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t  
A:Reference number: A90198; MUID:77021516  
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residu  
A:Accession: A90198  
A:Molecule type: protein  
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of t  
R;Alexander, A.; Steinmetz, M.; Barrilault, D.; Frangione, B.; Franklin, E.C.; Hood,  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deleti  
A:Reference number: A93915; MUID:82247835  
A:Contents: heavy chain disease protein Omm  
A:Accession: A93915  
A:Molecule type: mRNA  
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-  
A:Note: a carboxyl-terminal Lys is removed posttranslationally  
C:Comment: This sequence may represent an allelic form or another gamma chain subclass  
C:Genetics: The heavy chain disease protein Wis is shown.

```

Query Match      67.9%;   Score 1139;   DB 1;   Length 327;
Best Local Similarity 65.7%;   Pred. No. 7e-75;
Matches 226;   Conservative          9;   Mismatches          33;   Indels          76;   Gaps          4;

QY    24 SVAGEQAAGTAPCSRGGSSWSADLDKCMDCASCRRARPHSDFCLGCAAAP--PAPFRLLWRS 81
     1 |         ||||| : |         |         |         |         |         |
Db    2 STKGPSVFPLAPCSRSTS-----ESTAAALGCLVKDYFEPPEVTVSWNS 43
     QY  82 -----CDKTH----- 86
     Db  44 GALTSGVHTFPVLIQSSGLYSLSVVTPVSSSLGTKTTCNVDDHKNTKVDKRVESKYG 103
     QY  87 -TCPCCPAPEAEGAPSFLFPFKPDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGYE 145
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 104 PPCSCTAPEFLGPSVFLFPFKPDITLMISRTPEVTCVVVDVSOEDPEVQFNWYVDGYE 163
     QY 146 VHNNAKTKPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTIISKAKGP 205
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 164 VHNNAKTKPREQNSTYRVSVLTVLHODWLNGKEYCKVSNKGLPSSIETKISKAKGP 223
     QY 206 REPOVYTLPRESREMTKNQVSLTCLVKGFGFYSDIAVENESGNQPENNYKTTTPVLDSGGS 265
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 224 REPOVYTLPPEQEEMTKNQVSLTCLVKGFGFYSDIAVENESGNQPENNYKTTTPVLDSGGS 283
     QY 266 FFLYSKLTVDKSRMQOGNWFSCSVMHAEALHHNYTKSLSLSPGK 309
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 284 FFLYSKLTVDKSRMOEGNWFSCSVMHAEALHHNYTKSLSLSGK 327

RESULT 10
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:reference number: I47158; MUID:95015845
A:Accession: I47159
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:CROSS-references: EMBL:U03779; NID:g433123; PIDN:AAA52217.1; PID:g433124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match      55.3%;   Score 928;   DB 2;   Length 328;
Best Local Similarity 52.7%;   Pred. No. 1.le-59;
Matches 183;   Conservative          36;   Mismatches          46;   Indels          82;   Gaps          7;

QY    26 AGEQAQGT---APCSRGSSWSADLDKCMDCASCRRARPHSDFCLGCAAAP--PAPFRLLWR 80
     1 |         ||||| : |         |         |         |         |         |
Db    1 APKTAPSYVPLAPCSRSTS-----GPNVALGCLASSYFPEPVTVTN 42
     QY  81 S-----CDKTH----- 86
     Db  43 SGALSSGVHTFPVLIQSSGLYSLSMWVTPASSLSKSYSVCNVNHPTATTKVDKRVGTKT 102
     QY  87 --TCPCCPAPEAEGAPSFLFPFKPDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGY 144
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 103 KPPECIPACESP-G-PSYFIFFPKPKDITLMISRTQVTCVVVDVSOENPEVQSFWYVDG 161
     QY 145 EVHNAKTKPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTIISKAKGO 204
           ||||| |         ||||| |         ||||| |         ||||| |         ||||| |
     Db 162 EVHTAQTRKPEQNSTYRVSVLPIQHODWLNGEFKCKVNNKDLPAPITRIISKAKGO 221
     QY 205 PREPOVYTLPRESREMTKNQVSLTCLVKGFGFYSDIAVENESNGO--PENNYKITPPVLDS 262

```

[illegible]



C:Accession: I47162  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130  
C:Genetics:  
A:Gene: IgG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 54.1%; Score 907; DB 2; Length 277;  
Best Local Similarity 72.1%; Pred. No. 2.9e-58;  
Matches 165; Conservative 29; Mismatches 31; Indels 4; Gaps 3;

QY 85 HTCPCP-APEAG-APSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 142  
DB 49 TKTKPCICACEGPFSAIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 108  
QY 143 GVEVHNAKTKEEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 202  
DB 109 GVEVHNAKTKEEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 168  
QY 203 GOPREPQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVL 260  
DB 169 GQTRPQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVL 228  
QY 261 DSDGFELYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 309  
DB 229 DSDGFELYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 277

RESULT 14  
I47158  
Ig gamma 1 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
A:Reference number: I47158  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47158  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03778; NID:9433121; PIDN:AAA52216.1; PID:9433122  
C:Genetics:  
A:Gene: IgG1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 53.3%; Score 895; DB 2; Length 328;  
Best Local Similarity 52.2%; Pred. No. 2.6e-57;  
Matches 181; Conservative 31; Mismatches 53; Indels 82; Gaps 7;

QY 26 AGEQAPGT---APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFRLWR 80  
DB 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTMTWN 42  
QY 81 S-----CDKTH----- 86  
DB 43 SGALTSGVHTFSPVLPQSLGSLSSMTVPASSLSKSYTCNVNHPATTKVDKRVGIHQ 102

QY 87 --TCPPCPAPEAGAPSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 144  
DB 103 PQTCPICGCEVAG-PSVFIAPPKPKDTLMISQTEVTCVVVDVSHEDPEVKFNWYVDG 161

QY 145 EVHNAKTKEEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAKQ 204  
DB 162 EVHTAETRPKEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAKQ 221  
QY 205 PREPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVLDS 262  
DB 222 SREPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVLDS 281  
QY 263 DGSFFLYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 309  
DB 282 DGSFFLYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 328

RESULT 15  
I47161  
Ig gamma 3 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
A:Reference number: I47161  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47161  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128  
C:Genetics:  
A:Gene: IgG3  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 53.2%; Score 893; DB 2; Length 328;  
Best Local Similarity 52.3%; Pred. No. 3.6e-57;  
Matches 181; Conservative 33; Mismatches 52; Indels 80; Gaps 8;

QY 26 AGEQAPGT---APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFRLWR 80  
DB 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTMTWN 42  
QY 81 S-----CDKTH----- 88  
DB 43 SGALTSGVHTFSPVLPQSLGSLSSMTVPASSLSKSYTCNVNHPATTKVDKRVGTKT 102

QY 89 -PPCP-APEAGA-PSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 145  
DB 103 KPPCPICGCEVAGPSVFIAPPKPKDTLMISQTEVTCVVVDVSHEDPEVKFNWYVDGVE 162  
QY 146 VHNATKPREQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAKQ 205  
DB 163 VHTAETRPKEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAKQ 222  
QY 206 REPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVLDS 263  
DB 223 REPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPVLDS 282

QY 264 GSFFLYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 309  
DB 283 GTFFLYSLTVKDSRWQGNVFCSCVMHEALHNYTKQSLSPGK 328

Search completed: March 11, 2002, 15:14:51  
Job time: 135 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:12:36 ; Search time 53.92 Seconds  
(without alignments)  
424,493 Million cell updates/sec

Title: US-09-742-454A-7

Perfect score: 1678

Sequence: 1 MARGSLRLRLVLGLWLA.....MHEALHHYTKSLSLSPGK 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description           |
|------------|-------|-------------|--------|----------|-----------------------|
| 1          | 1678  | 100.0       | 309    | AAU03500 | Human TWEAKR receptor |
| 2          | 1258  | 75.0        | 451    | AA14765  | Human soluble kit     |
| 3          | 1255  | 74.8        | 438    | AA14765  | Human soluble kit     |
| 4          | 1255  | 74.8        | 438    | AA14765  | Human soluble kit     |
| 5          | 1250  | 74.5        | 660    | AA14765  | Human soluble kit     |
| 6          | 1250  | 74.5        | 660    | AA14765  | Human soluble kit     |
| 7          | 1250  | 74.5        | 660    | AA14765  | Human soluble kit     |
| 8          | 1250  | 74.5        | 660    | AA14765  | Human soluble kit     |
| 9          | 1247  | 74.3        | 376    | AA14765  | Human soluble kit     |
| 10         | 1247  | 74.3        | 450    | AA14765  | Human soluble kit     |
| 11         | 1247  | 74.3        | 453    | AA14765  | Human soluble kit     |

|    |        |      |     |    |           |                   |
|----|--------|------|-----|----|-----------|-------------------|
| 12 | 1243   | 74.1 | 424 | 16 | AAW14764  | Human soluble kit |
| 13 | 1243   | 74.1 | 424 | 16 | AAW14765  | Human soluble kit |
| 14 | 1239.5 | 73.9 | 375 | 18 | AAW06683  | Human soluble kit |
| 15 | 1239   | 73.8 | 488 | 21 | AAW06683  | Human soluble kit |
| 16 | 1239   | 73.8 | 535 | 20 | AAW17414  | Human soluble kit |
| 17 | 1235.5 | 73.6 | 234 | 18 | AAW16686  | Human soluble kit |
| 18 | 1235   | 73.6 | 400 | 21 | AAW15123  | Human soluble kit |
| 19 | 1233.5 | 73.5 | 396 | 18 | AAW18574  | Human soluble kit |
| 20 | 1233.5 | 73.5 | 396 | 18 | AAW18575  | Human soluble kit |
| 21 | 1232   | 73.4 | 376 | 18 | AAW050287 | Human soluble kit |
| 22 | 1230.5 | 73.3 | 302 | 22 | AAW050287 | Human soluble kit |
| 23 | 1230.5 | 73.3 | 302 | 22 | AAW050287 | Human soluble kit |
| 24 | 1230.5 | 73.3 | 302 | 22 | AAW050287 | Human soluble kit |
| 25 | 1230.5 | 73.3 | 302 | 22 | AAW050287 | Human soluble kit |
| 26 | 1230   | 73.3 | 232 | 22 | AAW02642  | Human soluble kit |
| 27 | 1230   | 73.3 | 292 | 22 | AAW02646  | Human soluble kit |
| 28 | 1230   | 73.3 | 292 | 22 | AAW02647  | Human soluble kit |
| 29 | 1230   | 73.3 | 330 | 20 | AAW05153  | Human soluble kit |
| 30 | 1230   | 73.3 | 339 | 22 | AAW02648  | Human soluble kit |
| 31 | 1230   | 73.3 | 401 | 18 | AAW10537  | Human soluble kit |
| 32 | 1230   | 73.3 | 444 | 21 | AAW32263  | Human soluble kit |
| 33 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 34 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 35 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 36 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 37 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 38 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 39 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 40 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 41 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 42 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 43 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 44 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |
| 45 | 1230   | 73.3 | 451 | 20 | AAW50031  | Human soluble kit |

#### ALIGNMENTS

#### RESULT 1

AAU03500  
ID AAU03500 standard; Protein; 309 AA.

AC AAU03500;

XX 26-SEP-2001 (first entry)

XX Human TWEAK receptor-Fc (TWEAKR-Fc) fusion polypeptide.

DE TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;  
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
KW corneal graft neovascularisation; psoriasis; metastatic condition;  
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
KW peripheral atherosclerosis; human IgG1; TWEAKR-Fc; fusion protein.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers  
FT Peptide 1..27 /note= "Signal peptide"

FT Protein 28..309 /note= "Mature human TWEAKR-Fc fusion protein."

FT Domain 28..79 /note= "Specifically referred to in Claim 11"

FT Region 80..81 /note= "From TWEAKR extracellular domain. Specifically referred to in Claims 4 and 10"





```
XX PA (IMMV ) IMMUNEX CORP.
XX PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX DR WPI; 1998-427099/36.
XX DR N-PSDB; AAV32636.
XX PT Purified polypeptide OX-40 ligands - for co-stimulation of T-cell
XX PT production and binding assays for OX-40 and homologues
XX PS Example 2; Col 37-40; 26pp; English.
XX CC The present sequence represents the OX40/Fc fusion protein that
XX CC contains the extracellular domain of mouse OX40 fused to the mutated
XX CC Fc region of the human IgG1 antibody. The fusion protein was used
XX CC for detecting cDNA clones encoding an OX40 ligand. The invention
XX CC claims for a murine OX40-L cytokine (AAW48975) that binds to the murine
XX CC T cell antigen, OX40. The OX40-L protein is claimed to be useful for
XX CC co-stimulation of T-cell production and in binding assays for
XX CC detecting OX40 or its homologues. The OX40-L protein is also claimed
XX CC to generate a TH-2 immune response.
XX SQ Sequence 438 AA;
Query Match 74.8%; Score 1255; DB 19; Length 438;
Best Local Similarity 65.8%; Pred. No. 2e-87;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
QY 35 PCSRG-SSWADLKKMDCAS-----CRAP-----HSDFCLG- 66
DB 62 pcegtfyneavnydtkqctqcnhrsgselkqncptqdtvcrpqtqrgdsgyklgv 121
QY 67 -CAAPAPF-----RLW----- 79
DB 122 dcvcpcppghspgnnqackpwtncitlsgkqtrhpasdsldavcedrslallilwetqrpt 181
QY 80 -----RSCDKHTHTCPPCPAPEGAPSVFLFPPPKDT 112
DB 182 frptvgsttwrtseipstptlveprscdkthtccpapeagapsvflfppkpkdt 241
QY 113 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 172
DB 242 lmsrtpevtcvvdvshedpevkfnwydgvvghnahtkpreeqynstyrsvsvltvlh 301
QY 173 ODWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVK 232
DB 302 qdwlngkeykckvsnkalpapiektiskakgqprepqvtytlppsrdeitknqsvltclvk 361
QY 233 GFVPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSVSMHE 292
DB 362 gfypsdiavewesngqpennnykttppvldsdgsfflyskltvdksrwqqgnvfscvmhe 421
QY 293 ALHNHYTQKSLSLSPGK 309
DB 422 alhnhytqkslsispgk 438
RESULT 5
AAV13463
ID AAV13463 standard; Protein; 660 AA.
XX AC AAV13463;
XX DT 26-JUL-1999 (first entry)
XX CC Ectromelia A39R semaphorin polypeptide.
XX KW VESPR; viral-encoded semaphorin protein receptor; semaphorin; IL-12;
XX KW inflammatory disease; viral infection; immune regulation; interleukin-12;
XX KW Th1 type immune response; cancer; autoimmune disease; A39.
```

```
OS Ectromelia virus.
XX PN WO9921997-A1.
XX PD 06-MAY-1999.
XX PF 28-OCT-1998; 98WO-US22879.
XX PR 28-OCT-1997; 97US-0112009.
XX PA (IMMV ) IMMUNEX CORP.
XX PI Comeau MR, Dubose RE, Johnson RS, Spriggs MK;
XX DR WPI; 1999-326704/27.
XX DR N-PSDB; AAX55565.
XX PT Viral-encoded protein receptors useful for treating inflammatory
XX PT diseases
XX PS Example 1; Page 63-66; 73pp; English.
XX CC The invention relates to a VESPR (viral-encoded semaphorin protein
XX CC receptor) polypeptide that can bind semaphorins. The VESPR polypeptides
XX CC are used to treat inflammatory diseases; when immobilized, to purify and
XX CC separate semaphorins or cells that express them. They are useful as
XX CC reagents for detecting, or measuring biological activity of, optionally
XX CC modified semaphorins; as carriers for delivering diagnostic and
XX CC therapeutic agents to semaphorin-positive cells; for studying the role of
XX CC VESPR in viral infection and immune regulation. When administered, in
XX CC combination with vaccination, together with a semaphorin, VESPR may
XX CC stimulate production of interleukin-12 (IL-12) which acts as an adjuvant
XX CC to induce a more persistent cellular immune response (of Th1 type),
XX CC including a curative response against aggressive, micrometastatic
XX CC cancers. Antibodies against VESPR can be used to treat autoimmune
XX CC diseases where an inflammatory response follows presentation of self-
XX CC antigens by T cells. Sense and antisense fragments of nucleic acid
XX CC encoding VESPR can be used to block expression of the polypeptide. The
XX CC present sequence represents an eromelia A39 semaphorin protein. The
XX CC native VESPR polypeptide is isolated from the human cells expressing the
XX CC receptor by using an Ectromelia virus A39 semaphorin/fc Fusion protein.
XX SQ Sequence 660 AA;
Query Match 74.5%; Score 1250; DB 20; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.8e-87;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 RSCDKHTHTCPPCPAPEGAPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 34 rscdkthtccpapeagapsvflfppkdtlmsrtpevtcvvdvshedpevkfnw 93
QY 140 YDGVGEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
DB 94 ydgvgevhnahtkpreeqynstyrsvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 200 KAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 259
DB 154 kagqprepqvtytlppsrdeemtknqsvltclvkgyfypsdlavewesngqpennnykttppv 213
QY 260 LDSGGSFFLYSKLTVDKSRWQQGNVFSVSMHEALHNHYTQKSLSLSPGK 309
DB 214 ldsdgsfflyskltvdksrwqqgnvfscsvmhealhnhytqkslsispgk 263
RESULT 6
AAB28523
ID AAB28523 standard; Protein; 660 AA.
XX AC AAB28523;
XX DT 07-FEB-2001 (first entry)
```

XX DE Ectromelia A39R semaphorin.  
XX DE Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;  
KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;  
KW inflammation; immune regulation; viral infection.  
XX OS Ectromelia virus.  
XX US6130068-A.  
XX PD 10-OCT-2000.  
XX PF 28-OCT-1998; 98US-0181706.  
XX PR 26-OCT-1998; 98US-0112009.  
XX PA (IMMV ) IMMUNEX CORP.  
XX PI Comeau MR, Johnson RS, Spriggs MK, Dubose RF;  
XX WPI; 2000-646753/62.  
XX DR N-PSDB; AAC63726.  
XX PT New Viral Encoded Semaphorin Protein Receptor DNA and polypeptides,  
XX PT useful for treating inflammation or inflammatory diseases, e.g.  
XX PT rheumatoid arthritis -  
XX PS Example 1; Column 55-60; 32pp; English.  
XX CC The present sequence is provided in an invention relating to novel  
XX CC semaphorin receptor polypeptides designated Viral Encoded Semaphorin  
XX CC Protein Receptor (VESPR). The DNA encoding the VESPR is useful for  
XX CC constructing expression vectors that produce the VESPR polypeptide. The  
XX CC VESPR DNA and polypeptide are useful for treating inflammation and  
XX CC inflammatory diseases, e.g. rheumatoid arthritis, a disease associated  
XX CC with chronic inflammation of the synovial tissue. The VESPR polypeptide  
XX CC is useful for measuring the biological activity of semaphorin proteins  
XX CC in terms of their binding affinity for VESPR. It is also useful as a  
XX CC research tool for studying the role that the receptor, in conjunction  
XX CC with semaphorins, may play in immune regulation and viral infection. The  
XX CC VESPR polypeptides are also useful as reagents that may be employed by  
XX CC those conducting quality assurance studies, e.g. to monitor shelf life  
XX CC and stability of semaphorin protein under different conditions.  
XX SQ Sequence 660 AA;  
Query Match 74.5%; Score 1250; DB 21; Length 660;  
Best Local Similarity 100.0%; Pred. No. 7.8e-87;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 80 RSCDKTHTCPCPAPEAGAPSVFLPPPKPDTLMISRPEVTCVVVDVSHEDPEVKFNW 139  
Db 34 rscdkthtccpcpapeagapsvflfppkpkdtlmsrptevtcvvvdvshedpevkfnw 93  
QY 140 YVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 199  
Db 94 yvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektis 153  
QY 200 KAKGQPREQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 259  
Db 154 kkgqprepqvtytlppsreemtknqvs l tclvkgfy psdiavwesngopennykttppv 213  
QY 260 LQSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309  
Db 214 lqsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispkg 263  
RESULT 7  
AAB70132  
ID AAB70132 standard; Protein; 660 AA.  
XX

AC AAB70132;  
XX 21-MAY-2001 (first entry)  
XX DE Ectromelia A39R semaphorin.  
XX DE Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;  
KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;  
KW inflammation; immune regulation; viral infection.  
XX OS Ectromelia virus.  
XX US6187909-B1.  
XX PD 13-FEB-2001.  
XX PF 10-DEC-1999; 99US-0459066.  
XX PR 26-OCT-1998; 98US-0112009.  
XX PR 28-OCT-1998; 98US-0181706.  
XX PR 28-OCT-1997; 97US-0958598.  
XX PA (IMMV ) IMMUNEX CORP.  
XX PI Spriggs MK, Comeau MR, Dubose RF, Johnson RS;  
XX WPI; 2001-256141/26.  
XX DR N-PSDB; AAF76950.  
XX PT New viral encoded semaphorin protein receptor polypeptides useful for  
XX PT treating inflammation or diseases associated with pro-inflammatory  
XX PT activity of a semaphorin ligand -  
XX PS Example 1; Column 55-60; 32pp; English.  
XX CC The DNA encoding the present sequence was used to prepare an Ectromelia  
XX CC semaphorin/Fc fusion protein. The fusion protein was used in the  
XX CC isolation of a human viral encoded semaphorin protein receptor (VESPR).  
XX CC The VESPR polypeptides are useful for treating inflammation, or diseases  
XX CC associated with pro-inflammatory activity of a semaphorin ligand. The  
XX CC polypeptides are also useful as carriers for delivering diagnostic or  
XX CC therapeutic agents to cells expressing semaphorins.  
XX SQ Sequence 660 AA;  
Query Match 74.5%; Score 1250; DB 22; Length 660;  
Best Local Similarity 100.0%; Pred. No. 7.8e-87;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 80 RSCDKTHTCPCPAPEAGAPSVFLPPPKPDTLMISRPEVTCVVVDVSHEDPEVKFNW 139  
Db 34 rscdkthtccpcpapeagapsvflfppkpkdtlmsrptevtcvvvdvshedpevkfnw 93  
QY 140 YVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 199  
Db 94 yvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektis 153  
QY 200 KAKGQPREQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 259  
Db 154 kkgqprepqvtytlppsreemtknqvs l tclvkgfy psdiavwesngopennykttppv 213  
QY 260 LQSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309  
Db 214 lqsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispkg 263  
RESULT 8  
AAB31694  
ID AAB31694 standard; Protein; 660 AA.  
XX  
AC AAB31694;  
XX

|        |   |   |     |
|--------|---|---|-----|
| DT     | 30-APR-2001                                   | (first entry)   |     |
| XX     |   |   |     |
| DE     |   | Amino acid sequence of the Ectromelia virus A39R semaphorin.              |     |
| DE     |   |   |     |
| KW     |   | Viral encoded semaphorin protein receptor; VESPR; semaphorin; interferon; |     |
| KW     |   | viral semaphorin ligand; Ectromelia A39R; Staphylococcus aureus; IL-12;   |     |
| KW     |   | interleukin-12; dendritic cell; cytokine production; immunomodulator;     |     |
| KW     |   | proinflammatory cytokine; Th1 cell differentiation; inflammation;         |     |
| KW     |   | inflammatory disease; tumour.   |     |
| XX     |   |   |     |
| OS     |   | Ectromelia virus.   |     |
| PN     | US6174689-B1.                                 |   |     |
| PN     |   |   |     |
| PD     | 16-JAN-2001.                                  |   |     |
| XX     |   |   |     |
| PF     | 10-DEC-1999;                                  | 99US-0458791.   |     |
| XX     |   |   |     |
| PR     | 26-OCT-1998;                                  | 98US-0112009.   |     |
| PR     | 28-OCT-1998;                                  | 98US-0181706.   |     |
| PR     | 27-OCT-1997;                                  | 97US-0958598.   |     |
| XX     |   |   |     |
| PA     | (IMMV ) IMMUNEX CORP.                         |   |     |
| XX     |   |   |     |
| PI     | Spriggs MK, Comeau MR, Dubose RF, Johnson RS; |   |     |
| XX     |   |   |     |
| DR     | WPI: 2001-167789/17.                          |   |     |
| DR     | N-PSDB; AAF25254.                             |   |     |
| XX     |   |   |     |
| PT     |   | Screening for binding to viral encoded semaphorin protein receptor,       |     |
| PT     |   | comprises contacting a mixture containing semaphorin, or cells that       |     |
| PT     |   | express semaphorin with protein, and detecting binding to protein         |     |
| XX     |   |   |     |
| PS     | Example 1; Column 55-60; 32pp; English.       |   |     |
| XX     |   |   |     |
| CC     |   | The present sequenc represents an Ectromelia virus A39R semaphorin        |     |
| CC     |   | ligand. The protein was used to isolate a viral encoded semaphorin        |     |
| CC     |   | protein receptor (VESPR) polypeptide from human cells. Semaphorins        |     |
| CC     |   | interact with their membrane bound receptors to synergise with interferon |     |
| CC     |   | and Staphylococcus aureus (type C) in the production of interleukin-12    |     |
| CC     |   | (IL-12) from dendritic cells. VESPR can therefore be used to induce       |     |
| CC     |   | IL-12 production which in turn promotes cytokine production. IL-12 is     |     |
| CC     |   | a proinflammatory cytokine and an immunomodulator. A soluble VESPR can    |     |
| CC     |   | be used to antagonise IL-12 and downregulate Th1 cell differentiation.    |     |
| CC     |   | VESPR is therefore useful for treating inflammation and inflammatory      |     |
| CC     |   | diseases. VESPR ligands may be used to induce an immune response against  |     |
| CC     |   | aggressive tumours.   |     |
| XX     |   |   |     |
| SQ     | Sequence                                      | 660 AA;   |     |
|        | Query Match                                   | 74.5%; Score 1250; DB 22; Length 660;                                     |     |
|        | Best Local Similarity                         | 100.0%; Pred. No. 7.8e-87;  |     |
|        | Matches 230; Conservative                     | 0; Mismatches 0; Indels 0; Gaps 0   |     |
| QY     | 80  | RSCDKTKHTCCPCAPAEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW               | 139 |
| Db     | 34  | rscdktkhtccpcapaeagapsvflfppkpkdtlmisrtpevtcvvvdvshdepevkfnw              | 93  |
| QY     | 140   | YVDGVEVHNAKTKPREQYNSTRYVVSUVTLVLHQDLNGKEYCKVSNKALPAPIEKTIS                | 199 |
| Db     | 94  | yvdgvevhnaktckpreeqynstryrvsvvltvlhqdwlngkeyckvsnkalpapiektis             | 153 |
| QY     | 200   | KAKGQPREQVYTLPPSREEMTKNQVSLTCLVKGYGPPSDIAVEWESNGQPENNYKTPPV               | 259 |
| Db     | 154   | kakgqpreqvyltppsreemtknqvsitclvkgyfypsdiavewesngqpennyktppv               | 213 |
| QY     | 260   | LDSGSGFELYSLKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK                          | 309 |
| Db     | 214   | ldsgsgfflyslkltvdksrwqgnvfscvmealhnhytqkslsispgk                          | 263 |
| RESULT | 9   |   |     |









CC They are also useful in screening assays for (ant)agonists of HER3  
CC and HER4, and bind to cells that express HER4 e.g. MDA-MB-453 human  
CC breast cancer cells, inducing expression of intercellular adhesion  
CC molecule-1.  
XX  
SQ Sequence 375 AA;

Query Match 73.9%; Score 1239.5; DB 18; Length 375;  
Best Local Similarity 93.9%; Pred. No. 2.5e-86;  
Matches 229; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 66 GCAAPAPRLWRSCDKTHTPCPCPAEAGAPSVFLPPPKDITLMSRTPEVTCV 125  
Db 137 gfsgdpep-----kscdkthtccpapefegapsvflfppkdtlmsrtpevtcv 191  
QY 126 VDVSHEDEPKFNNYDGVVHNAKTPREEQYNSTYRVSVTLVHQDLNKGKCKV 185  
Db 192 vdvshdedevkfnwydgvvhnaktkpreeqynstyrsvsvtlvhdwlngkyckv 251  
QY 186 SNKALPAPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLCLVKGFYVPSDIAVEWES 245  
Db 252 snkalpapiiektiskagoprepqvytlppsrdeitknqvsitclvkgyfypsdiawes 311  
QY 246 NGOPENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSL 305  
Db 312 ngopennyktpvldsdgsfflyskltvdksrwqgnvfscsvmhealhnyhtqkls 371  
QY 306 SPQK 309  
Db 372 spgk 375

RESULT 15  
AAV97175  
ID RAY97175 standard; Protein; 488 AA.  
XX  
AC AAV97175;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE Human FGF-RI Extracellular domain-Ig Fc fusion protein 6.  
XX  
KW FGF-R; fibroblast growth factor receptor; extracellular domain; IgG1;  
KW immunoglobulin; G1; oligomerization domain; Fc region; fusion protein;  
KW inhibitor; dimer; antagonist; cytostatic; anti-diabetic; vulnerary;  
KW ophthalmological; anti-proliferative.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1..21  
FT /label= FGF-RI\_signal\_peptide  
FT Domain 22..257  
FT /label= FGF-RI\_extracellular\_domain  
FT /note= "the Ig I segment and acid box are deleted"  
FT Domain 59..111  
FT /label= Ig\_II\_segment  
FT Domain 157..222  
FT /label= Ig\_III\_segment  
FT Region 258..488  
FT /label= Human\_IgG1\_Fc\_region  
FT /note= "Contains hinge region and domains CH2 and CH3"  
FT Misc-difference 276  
FT /label= L276E  
FT /note= "this mutation decreases the affinity of the Fc  
FT portion for Fc receptors"  
FT Misc-difference 376  
FT /label= p376S  
FT /note= "this mutation decreases the affinity of the Fc  
FT portion for complement"  
XX  
PN WC2000046380-A2.

XX 10-AUG-2000.  
XX  
XX  
PF 07-FEB-2000; 2000WO-US03166.  
XX  
PR 08-FEB-1999; 99US-0119002.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Kavanaugh WM, Ballinger M;  
XX  
DR WPI; 2000-514961/46.  
XX  
DR N-PSDB; AAA52132.  
XX  
PT New polypeptide comprising a fibroblast growth factor receptor  
PT extracellular domain fused to a heterologous oligomerization domain for  
PT treating FGF-, angiogenesis-, or FGF receptor-mediated disorders  
XX  
XX Claim 14; Page 68-69; 70pp; English.  
XX  
CC Novel fusion protein constructs comprise a fibroblast growth factor (FGF)  
CC receptor (FGF-R) extracellular domain (ECD) lacking the immunoglobulin  
CC (Ig) I segment fused to a heterologous oligomerization domain that  
CC comprises an immunoglobulin Fc region, hinge region, CH1, CH2, CH3 or CH4  
CC region, or light chain of an immunoglobulin molecule, or a peptide with a  
CC leucine zipper motif. The Ig I segment is not necessary for binding of  
CC acidic FGF and basic FGF (bFGF). The Ig I deletion further increases the  
CC affinity for aFGF and heparin, protects the core of the molecule from  
CC proteolysis, and abrogates the heparin requirement for aFGF binding. The  
CC new fusion polypeptides are better FGF inhibitors than FGF-R monomer  
CC proteins. The FGF-R Ig Fc fusion dimers are active as FGF antagonists at  
CC subnanomolar concentrations and were 20-fold more potent than the FGF-R  
CC monomer protein as competitors of bFGF binding to immobilized FGF-Rs. The  
CC fusion constructs are useful to treat FGF-, angiogenesis- or  
CC FGF-R-mediated disorders, such as tumorigenesis (e.g. bladder, breast,  
CC lung, rectal, testis and cervical tumours), neovascularization (e.g.  
CC diabetic retinopathy, neovascular glaucoma, wound healing and corneal  
CC scarring) and hyper-proliferation of vascular smooth muscle cells (e.g.  
CC postangioplasty and postatherectomy restenosis).  
XX  
SQ Sequence 488 AA;

Query Match 73.8%; Score 1239; DB 21; Length 488;  
Best Local Similarity 81.1%; Pred. No. 3.7e-86;  
Matches 240; Conservative 9; Mismatches 23; Indels 24; Gaps 4;

QY 22 LRSVAGEQAPGTAPCSRSG-----SSWSADLDKCMDCASCRARPHSDCLGCAAPPA 73  
Db 209 lrnvstfeda-geytcclagnsigishhsawitvie-----aleerp-----avmts 252  
QY 74 PFRLLWRSCDKTHTPCPCPAEAGAPSVFLPPPKDITLMSRTPEVTCVVDVSHEDP 133  
Db 253 plylepkscdkthtccpapelegpsvflfppkdtlmsrtpevtcvvvdvshedp 312  
QY 134 EVAFNFWVDCGEVHNNAKTPREEQYNSTYRVSVTLVHQDLNKGKCKVSNKALPAP 193  
Db 313 evkfnwvvdgvevhnaktkpreeqynstyrsvsvtlvhdwlngkyckvsnkalpas 372  
QY 194 IEKTISKAGOPREPQVYTLPPSREEMTKNOVSLCLVKGFYVPSDIAVEWESNGQPENNY 253  
Db 373 iektiskagoprepqvytlppsrdeitknqvsitclvkgyfypsdiawesngqpenny 432  
QY 254 KTTTPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPGK 309  
Db 433 ktppvldsdgsfflyskltvdksrwqgnvfscsvmhealhnyhtqkls 488

Search completed: March 11, 2002, 15:14:13  
Job time: 97 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:53 ; Search time 31.92 Seconds  
(without alignments)  
672.971 Million cell updates/sec

Title: us-09-742-454a-7\_COPY\_28\_309

Perfect score: 1551

Sequence: 1 EQAPGTAPCGSSGWSADLD.....MHEALHNHYTKSLSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 1230   | 79.3        | 374    | 2     | S69339      |
| 2          | 1224   | 78.9        | 255    | 4     | S31866      |
| 3          | 1224   | 78.9        | 330    | 1     | GHU         |
| 4          | 1171.5 | 75.5        | 377    | 2     | A23511      |
| 5          | 1171   | 75.5        | 234    | 2     | PT0207      |
| 6          | 1169.5 | 75.4        | 377    | 2     | A60764      |
| 7          | 1158.5 | 74.7        | 326    | 1     | G2HU        |
| 8          | 1146.5 | 73.9        | 289    | 1     | G3HUI       |
| 9          | 1138   | 73.4        | 327    | 1     | G4HU        |
| 10         | 927    | 59.8        | 328    | 2     | I47159      |
| 11         | 923    | 59.5        | 328    | 2     | I47160      |
| 12         | 910.5  | 58.7        | 323    | 1     | GHRB        |
| 13         | 907    | 58.5        | 277    | 2     | I47162      |
| 14         | 894    | 57.6        | 328    | 2     | I47158      |
| 15         | 892    | 57.5        | 328    | 2     | I47161      |
| 16         | 876    | 56.5        | 329    | 1     | G2GP        |
| 17         | 859.5  | 55.4        | 470    | 2     | S22080      |
| 18         | 842    | 54.3        | 472    | 2     | S31459      |
| 19         | 840    | 54.2        | 308    | 2     | C30554      |
| 20         | 840    | 54.2        | 444    | 2     | PC4436      |
| 21         | 838    | 54.0        | 326    | 2     | P50017      |
| 22         | 836    | 53.9        | 329    | 1     | G3MSC       |
| 23         | 833    | 53.7        | 324    | 1     | G1MS        |
| 24         | 828    | 53.4        | 393    | 1     | G1MSM       |
| 25         | 825    | 53.2        | 333    | 2     | P50018      |
| 26         | 825    | 53.2        | 398    | 1     | G3MSM       |
| 27         | 817    | 52.7        | 329    | 2     | S00847      |
| 28         | 804.5  | 51.9        | 322    | 2     | P50019      |
| 29         | 799.5  | 51.5        | 330    | 1     | G2MSA       |

30 799.5 51.5 469 2 S37483  
31 794.5 51.2 399 1 G2MSAM  
32 794 51.2 335 1 G2MSAB  
33 786.5 50.7 474 1 G2MS11  
34 786 50.7 446 2 S40295  
35 781.5 50.4 405 1 G2MSBM  
36 776.5 50.1 327 2 S06611  
37 762.5 49.2 475 2 S01321  
38 708 45.6 180 2 I46732  
39 583.5 37.6 249 2 S69340  
40 577 37.2 152 2 S14236  
41 568.5 36.7 218 2 A36040  
42 410 26.4 572 2 B46529  
43 367 23.7 453 2 S37768  
44 366.5 23.6 388 1 EHMS  
45 365 23.5 391 1 MHUBT

#### ALIGNMENTS

##### RESULT 1

S69339

Ig heavy chain V region precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000

C:Accession: S69339; S72664

R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.

Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.

A:Reference number: S69339; MUID:95262687

A:Accession: S69339

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

R:Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994

A:Reference number: S72664

A:Accession: S72664

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140,'C',142-374 <KH2>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 79.3%; Score 1230; DB 2; Length 374;  
Best Local Similarity 98.3%; Pred. No. 8.1e-82;  
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDTHTCPPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

:|||||

Db 145 KSCDTHTCPPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 204

QY 113 YVDGVEVHNATKPREQVNSTYRVVSVTLVHQQWLNKCKVSKNKPAPIEKTIS 172

:|||||

Db 205 YVDGVEVHNATKPREQVNSTYRVVSVTLVHQQWLNKCKVSKNKPAPIEKTIS 264

QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 232

:|||||

Db 265 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 324

QY 233 LQSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282

:|||||

Db 325 LQSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 374

##### RESULT 2

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli









|   |              |  |                |             |
|---|--------------|--|----------------|-------------|
| Query Match   | 73.4%;       | Score 1138;  | DB 1;          | Length 327; |
| Best Local Similarity   | 67.1%;       | Pred. No. 3  | 1e-75;         |             |
| Matches 224;  | Conservative | 9;   | Mismatches 25; | Indels 76;  |
| Gaps  |              |  |                |             |
| QY  | 7            | APCSRGSSWSADLDKCMDCASCRRPHSDFCIGCAAAP--PAPFRLLRWS-----       | 54             |             |
| DB  | 12           | APCSRSTS-----ESTAAALGCLVKDYFPPEVTVSWNSGALTSQVHTF             | 53             |             |
| QY  | 55           | -----CDKTH-----TCPPCPAPE                                     | 68             |             |
| DB  | 54           | PVLQSSGLYSLSVVTVPPSSSLGTQTYCTNDVHDKTSNTKVDKRRVSKYGGPCPCPAPPE | 113            |             |
| QY  | 69           | AEGAPSVFLPPPKPDKDTLMSRTPETVCVVVDVSHEDPEVKENWYVDGEVHNAKTKPKE  | 128            |             |
| DB  | 114          | FLGGPSVFLPPPKPDKDTLMSRTPETVCVVVDVSHEDPEVQFNWYVDGEVHNAKTKPKE  | 173            |             |
| QY  | 129          | EQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPVVTLPP  | 188            |             |
| DB  | 174          | EQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTIKAKGQPREPVVTLPP   | 233            |             |
| QY  | 189          | SREMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVD   | 248            |             |
| DB  | 234          | SOEEMTKNQVSLCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVD  | 293            |             |
| QY  | 249          | KSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK                            | 282            |             |
| DB  | 294          | KSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK                            | 327            |             |
| RESULT  | 10           |  |                |             |
| I47159  |              |  |                |             |
| Ig gamma 2a chain constant region - pig (fragment)                          |              |  |                |             |
| C:Species: Sus scrofa domestica (domestic pig)                              |              |  |                |             |
| C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000 |              |  |                |             |
| C:Accession: I47159   |              |  |                |             |
| R:KacsKovics, I.; Sun, J.; Butler, J. E.                                    |              |  |                |             |
| J. Immunol. 153; 3565-3573, 1994  |              |  |                |             |
| A:title: Five putative subclasses of swine IgG identified from the cDNA se  |              |  |                |             |
| A:Reference number: I47158; MUID:95015845                                   |              |  |                |             |
| A:Accession: I47159   |              |  |                |             |
| A:Status: preliminary; translated from GB/EMBL/DBJ                          |              |  |                |             |
| A:Molecule type: mRNA   |              |  |                |             |
| A:Residues: 1-328 <KAC>   |              |  |                |             |
| A:Cross-references: EMBL:U03779; NID:g433123; PID:AAA52217.1; PID:g433124   |              |  |                |             |
| C:Genetics:   |              |  |                |             |
| A:Gene: IgG2a   |              |  |                |             |
| C:Superfamily: Immunoglobulin C region; immunoglobulin homology             |              |  |                |             |
| F:133-202/Domain: Immunoglobulin homology <IMM>                             |              |  |                |             |
| Query Match   | 59.8%;       | Score 927;   | DB 2;          | Length 328; |
| Best Local Similarity   | 52.7%;       | Pred. No. 5.8e-60;   |                |             |
| Matches 183;  | Conservative | 35;  | Mismatches 43; | Indels 86;  |
| Gaps  |              |  |                |             |
| QY  | 3            | APGT-----APCSRGSSWSADLDKCMDCASCRRPHSDFCIGCAAAP--PAPFRLLR     | 53             |             |
| DB  | 1            | APTAPSVYPLAPCSRDTs-----GPNVALGCLASSYFPEPVTVTN                | 42             |             |
| QY  | 54           | S-----CDKTH-----   | 59             |             |
| DB  | 43           | SGALSGCVHTFPVLPQSGLYSLSSMVTVPASSLSKSYCTCNVHPATTTKVDKRVGKT    | 102            |             |
| QY  | 60           | --TCPPCPAPEAGASVFLFPPKPKDTLMSRTPETVCVVVDVSHEDPEVKFNWYVDG     | 117            |             |
| DB  | 103          | KPPCPTCAPCESPG--PSVFFFPKPKDTLMSRTPQVTCVVVDVSENPEVQFSWYVDG    | 161            |             |
| QY  | 118          | EVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG   | 177            |             |
| DB  | 162          | EVAHTAPRKPEQFNSTYRVVSVLPIQHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG   | 221            |             |
| QY  | 178          | PREPOVYTLPPSREMTKNQVSLCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLD    | 235            |             |

||||| 222 TREPVYTLPPHAEELSRKSVITCLVIGYPPDIDVEMQNGQPEEGNYRTTPQODV 281  
QY 236 DGSFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTKQSLSPGK 282  
Db 282 DGTLYLKSFKSVDKASWQGGGIFQCAVMHEALHNHYTKQSLSPGK 328  
RESULT 11  
I47160  
Ig gamma 2b chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47160  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47160  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126  
C:Genetics:  
A:Gene: IGG2b  
C:Superfamily: immunoglobulin C region: immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>  
Query Match 59.5%; Score 923; DB 2; Length 328;  
Best Local Similarity 52.4%; Pred. No. 1.1e-59;  
Matches 182; Conservative 35; Mismatches 44; Indels 86; Gaps 7;  
QY 3 APGT-----APCSRGSSWSADLCKMDCASCARPHSDFCGCAAP--PAPFLLWR 53  
Db 1 APKTAFLVPLAPGCRDTS-----GPNVALGCLASSYFPEPVTWN 42  
QY 54 S-----CDKTH----- 59  
Db 43 SGALTSGVHTFPSPVLPQSLGLYSSMTVVPASSLSKSYCNVNHPPATTKVDRVGTCT 102  
QY 60 --TCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGV 117  
Db 103 KPCCIPCAPCESPG-PSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGV 161  
QY 118 EVHNAKTPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKG 177  
Db 162 EVHTAQTFRPKEQFNSTYRVSVLTPLIQHODWLNGKEYCKVSNKALPAPIEKTISKAKG 221  
QY 178 PREPVYTLPPSRREMTKNQVSLTCLVKGFPSPDIKAVESWNGQ--PENNYKTTTPPVLD 235  
Db 222 TREPVYTLPPHAEELSRKSVITCLVIGYPPDIDVEMQNGQPEEGNYRTTPQODV 281  
QY 236 DGSFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTKQSLSPGK 282  
Db 282 DGTLYLKSFKSVDKASWQGGGIFQCAVMHEALHNHYTKQSLSPGK 328  
RESULT 12  
GHRB  
Ig gamma chain C region - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 16, 387-397, 1983  
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplo  
A:Reference number: A91749; MUID:84030930  
A:Accession: A91749  
A:Molecule type: mRNA  
A:Residues: 1-323 <BER>  
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr  
R:Pratt, D.M.; Mole, L.E.

Biochem. J. 151, 337-349, 1975  
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog  
A:Reference number: A90290; MUID:76135469  
A:Accession: A90290  
A:Molecule type: protein  
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy ch  
A:Reference number: A93928; MUID:83299917  
A:Accession: A93928  
A:Molecule type: mRNA  
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112  
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma  
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul  
A:Reference number: A90245; MUID:70110015  
A:Accession: A90245  
A:Molecule type: protein  
A:Residues: 132-143, 'E', 145-161 <PRU>  
R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse  
A:Reference number: A94416  
A:Accession: A94416  
A:Molecule type: protein  
A:Residues: 129-131, 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '  
A:Note: this has the e15 allotypic marker, 185-Ala  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-82/Domain: immunoglobulin homology <IMI>  
F:130-199/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 58.7%; Score 910.5; DB 1; Length 323;  
Best Local Similarity 52.7%; Pred. No. 8.9e-59;  
Matches 178; Conservative 31; Mismatches 54; Indels 75; Gaps 6;  
QY 2 QAPGT---APCSRGSSWSADLCKMDCASCARPHSDFCGCAAP--PAPFLLWRS-- 54  
Db 4 KASVFFLAPC-----CGDTPSSVTTLGCLVKGYPVPTVWNSGT 45  
QY 55 -----CDKTH-----TC--PPC 64  
Db 46 LTNGVRTFPSPVROSSGLYSLSSVSVTSVSSQPVTCNVAHPATNTKVDKTVAPSTCSKPTC 105  
QY 65 PAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKT 124  
Db 106 PPPELLGGPSVFIAPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFTWYINNEGVRTARP 165  
QY 125 KPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPVY 184  
Db 166 PLREQNPNSTIRVVSTLPTTHQDWLNGKEYCKVSNKALPAPIEKTISKARGQPLEPKVY 225  
QY 185 TLPPSRREMTKNQVSLTCLVKGFPSPDIKAVESWNGQPEENNYKTTTPPVLDSDGFFLYSK 244  
Db 226 TMGPPREELSSRSVSLTCMNGFYPSDISVEWENKGAEDNYKTTTPAVLSDSGSYFLYNK 285  
QY 245 LTVDKSRWQGNVSCSVNHEALHNHYTKQSLSPGK 282  
Db 286 LSVPTSEWQRGDFTCSVMHEALHNHYTKQSLSPGK 323  
RESULT 13  
I47162  
Ig gamma 4 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000

C:Accession: I47162  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130  
C:Genetics:  
A:Gene: IgG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 58.5%; Score 907; DB 2; Length 277;  
Best Local Similarity 72.1%; Pred. No. 1.3e-58;  
Matches 165; Conservative 29; Mismatches 31; Indels 4; Gaps 3;  
QY 58 THTCP-AP-APS-APSVLP-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVD 115  
Db 49 TTKTK-PP-PC-ACEG-PSAF-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVD 108  
QY 116 GVEVHNAKTP-REOYNSTYRVSVLTVLHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAK 175  
Db 109 GVEVHTA-TRP-KEEQFNSTYRVSVLPIQHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAK 168  
QY 176 GQRPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 233  
Db 169 GQRPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 233  
QY 234 DSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 282  
Db 229 DSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 282

RESULT 14  
I47158  
Ig gamma 1 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47158  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47158  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03778; NID:9433121; PIDN:AAA52216.1; PID:9433122  
C:Genetics:  
A:Gene: IgG1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 57.6%; Score 894; DB 2; Length 328;  
Best Local Similarity 52.2%; Pred. No. 1.4e-57;  
Matches 181; Conservative 30; Mismatches 50; Indels 86; Gaps 7;  
QY 3 APGT-----APCSRGSSWSADLCKMDCASCRARPHSD-FC-GLCAAAP--PAPFLLMR 53  
Db 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTWTWN 42  
QY 54 S-----CDKTHC----- 59  
Db 43 SGALTSGVHTFSPVLP-PSGLYSLSSVMVTPASSLSKSYTCNVNHPATTKVDKRVGIHQ 102  
QY 60 --TCPP-AP-APS-APSVLP-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVDGV 117  
Db 103 PQTCTPC-PC-CEVAG-PSVFI-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVDGV 161

QY 118 EVHNAKTP-REOYNSTYRVSVLTVLHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAKGQ 177  
Db 162 EVHTA-TRP-KEEQFNSTYRVSVLPIQHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAKGQ 221  
QY 178 PREPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 235  
Db 222 SREPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 281  
QY 236 DSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 282  
Db 282 DSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 328  
RESULT 15  
I47161  
Ig gamma 3 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47161  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47161  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128  
C:Genetics:  
A:Gene: IgG3  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 57.5%; Score 892; DB 2; Length 328;  
Best Local Similarity 52.3%; Pred. No. 2e-57;  
Matches 181; Conservative 32; Mismatches 49; Indels 84; Gaps 8;  
QY 3 APGT-----APCSRGSSWSADLCKMDCASCRARPHSD-FC-GLCAAAP--PAPFLLMR 53  
Db 1 APKTAPSVYPLAPCGRDTS-----GPNVALGCLASSYFPEPVTWTWN 42  
QY 54 S-----CDKTHC----- 61  
Db 43 SGALTSGVHTFSPVLP-PSGLYSLSSVMVTPASSLSKSYTCNVNHPATTKVDKRVGTXT 102  
QY 62 -PPCP-AP-APS-APSVLP-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVDGV 118  
Db 103 KPPCPC-PC-CEVAG-PSVFI-PPKPKD-MSRTP-ETV-CTV-VDV-SHED-PEVK-ENYVDGV 162  
QY 119 VINAKTP-REOYNSTYRVSVLTVLHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAKGQ 178  
Db 163 VHTA-TRP-KEEQFNSTYRVSVLPIQHODWLNK-KEYCKVSKN-ALPAPI-ETKTSKAKGQ 222  
QY 179 REPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 236  
Db 223 REPQVYTLPP-REEMTKNOVSLTCLVKGYPSDIAV-ENESNGQ--PENNYKTTTPPV 282  
QY 237 GSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 282  
Db 283 GSDGFFLYSKLTVDKSRWQGNVFS-VMHEALHNHYTK-QLSLS-SPGK 328

Search completed: March 11, 2002, 15:14:53  
Job time: 137 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:52 ; Search time 19.78 seconds  
(without alignments)  
522.724 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_309

Perfect score: 1551

Sequence: 1 EQAPGAPCSRGSWSADLD.....MHEALHNHYTKSLSLSPGK 282

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description        |
|------------|--------|---------------|--------|-------|--------------------|
| 1          | 1224   | 78.9          | 330    | 1     | GC1_HUMAN          |
| 2          | 1158.5 | 74.7          | 326    | 1     | P01857 homo sapien |
| 3          | 1151.5 | 74.2          | 290    | 1     | GC2_HUMAN          |
| 4          | 1138   | 73.4          | 327    | 1     | P01860 homo sapien |
| 5          | 910.5  | 58.7          | 323    | 1     | P01861 homo sapien |
| 6          | 876    | 56.5          | 329    | 1     | P01870 oryctolagus |
| 7          | 838    | 54.0          | 326    | 1     | P01862 cavia porce |
| 8          | 836    | 53.9          | 329    | 1     | P20759 rattus norv |
| 9          | 833    | 53.7          | 324    | 1     | P01868 mus musculu |
| 10         | 828    | 53.4          | 333    | 1     | P01869 mus musculu |
| 11         | 825    | 53.2          | 333    | 1     | P20761 rattus norv |
| 12         | 825    | 53.2          | 398    | 1     | P03987 mus musculu |
| 13         | 817    | 52.7          | 329    | 1     | P20762 rattus norv |
| 14         | 804.5  | 51.9          | 322    | 1     | P20760 rattus norv |
| 15         | 799.5  | 51.5          | 330    | 1     | P01863 mus musculu |
| 16         | 794.5  | 51.2          | 399    | 1     | P01865 mus musculu |
| 17         | 794    | 51.2          | 335    | 1     | P01864 mus musculu |
| 18         | 786.5  | 50.7          | 336    | 1     | P01866 mus musculu |
| 19         | 781.5  | 50.4          | 405    | 1     | P01867 mus musculu |
| 20         | 367    | 23.7          | 454    | 1     | P01871 homo sapien |
| 21         | 366.5  | 23.6          | 421    | 1     | P06336 mus musculu |
| 22         | 365    | 23.5          | 391    | 1     | P04220 homo sapien |
| 23         | 360.5  | 23.2          | 429    | 1     | P01855 rattus norv |
| 24         | 353.5  | 22.8          | 428    | 1     | P01854 homo sapien |
| 25         | 349    | 22.5          | 455    | 1     | P01872 mus musculu |
| 26         | 344    | 22.2          | 458    | 1     | P03988 oryctolagus |
| 27         | 339    | 21.9          | 476    | 1     | P01873 mus musculu |
| 28         | 336    | 21.7          | 454    | 1     | P06337 mesocricetu |
| 29         | 334    | 21.5          | 479    | 1     | P04221 oryctolagus |
| 30         | 333    | 21.5          | 457    | 1     | P20768 canis muri  |
| 31         | 331.5  | 21.4          | 450    | 1     | P01874 canis fami  |
| 32         | 302.5  | 19.5          | 299    | 1     | P01879 oryctolagus |
| 33         | 290    | 18.7          | 438    | 1     | P23085 heterodontu |

|    |       |      |     |   |            |
|----|-------|------|-----|---|------------|
| 34 | 288.5 | 18.6 | 446 | 1 | MUC_CHICK  |
| 35 | 284.5 | 18.3 | 438 | 1 | HVCS_HETFR |
| 36 | 282   | 18.2 | 353 | 1 | ALC1_HUMAN |
| 37 | 279   | 18.0 | 353 | 1 | ALC1_GORGO |
| 38 | 275.5 | 17.8 | 461 | 1 | HVCM_HETFR |
| 39 | 274.5 | 17.7 | 340 | 1 | ALC2_HUMAN |
| 40 | 274   | 17.7 | 370 | 1 | HVC1_HETFR |
| 41 | 273.5 | 17.6 | 393 | 1 | HVC3_HETFR |
| 42 | 255   | 16.4 | 344 | 1 | ALC_MOUSE  |
| 43 | 252.5 | 16.3 | 481 | 1 | MUCM_ICTFU |
| 44 | 178.5 | 11.5 | 105 | 1 | LAC1_MOUSE |
| 45 | 173   | 11.2 | 104 | 1 | LAC2_RAT   |

#### ALIGNMENTS

| RESULT | GC1_HUMAN   | STANDARD; | PRT; | 330 AA. |
|--------|---|-----------|------|---------|
| ID     | GC1_HUMAN   |           |      |         |
| AC     | P01857;   |           |      |         |
| DT     | 21-JUL-1986 (Rel. 01, Created)  |           |      |         |
| DT     | 21-JUL-1986 (Rel. 01, Last sequence update)   |           |      |         |
| DT     | 15-JUL-1999 (Rel. 38, Last annotation update)   |           |      |         |
| DE     | IG GAMMA-1 CHAIN C REGION.  |           |      |         |
| GN     | IGHG1   |           |      |         |
| OS     | Homo sapiens (Human).   |           |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;   |           |      |         |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  |           |      |         |
| OX     | NCBI_TaxID=9606;  |           |      |         |
| RN     | [1]   |           |      |         |
| RP     | SEQUENCE FROM N.A.  |           |      |         |
| RX     | MEDLINE=82274238; PubMed=6287432;   |           |      |         |
| RA     | Ellison J.W., Berson B.J., Hood L.E.;   |           |      |         |
| RT     | "The nucleotide sequence of a human immunoglobulin C gamma gene."   |           |      |         |
| RL     | Nucleic Acids Res. 10:4071-4079(1982).  |           |      |         |
| RN     | [2]   |           |      |         |
| RP     | SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).   |           |      |         |
| RX     | MEDLINE=71064024; PubMed=5489771;   |           |      |         |
| RA     | Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,  |           |      |         |
| RT     | Waxdal M.J., Edelman G.M.;  |           |      |         |
| RL     | "The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."   |           |      |         |
| RN     | Biochemistry 9:3161-3170(1970).   |           |      |         |
| RP     | SEQUENCE OF 136-329 (EU).   |           |      |         |
| RX     | MEDLINE=71064025; PubMed=5530842;   |           |      |         |
| RA     | Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,   |           |      |         |
| RT     | Edelman G.M.;   |           |      |         |
| RL     | "The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."   |           |      |         |
| RN     | Biochemistry 9:3171-3181(1970).   |           |      |         |
| RP     | SEQUENCE (MYELOMA PROTEIN NIE).   |           |      |         |
| RX     | MEDLINE=77070269; PubMed=826475;  |           |      |         |
| RA     | Ponstingl H., Hilschmann N.;  |           |      |         |
| RT     | "The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure." |           |      |         |
| RL     | Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).   |           |      |         |
| RN     | [5]   |           |      |         |
| RP     | SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.  |           |      |         |
| RX     | MEDLINE=83289131; PubMed=6884994;   |           |      |         |
| RA     | Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;   |           |      |         |
| RT     | "Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."   |           |      |         |
| RL     | Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).   |           |      |         |
| RN     | [6]   |           |      |         |
| RP     | DISULFIDE BONDS.  |           |      |         |
| RX     | MEDLINE=71064027; PubMed=4923144;   |           |      |         |
| RA     | Gall W.E., Edelman G.M.;  |           |      |         |
| RT     | "The covalent structure of a human gamma G-immunoglobulin. X."  |           |      |         |

|        |             |
|--------|-------------|
| P01875 | gallus gall |
| P23087 | heterodontu |
| P01876 | homo sapien |
| P20758 | gorilla gor |
| P23088 | heterodontu |
| P01877 | homo sapien |
| P23084 | heterodontu |
| P23086 | heterodontu |
| P01878 | mus musculu |
| P23735 | ictalurus p |
| P01843 | mus musculu |
| P20767 | rattus norv |



|     |          |         |           |  |
|-----|----------|---------|-----------|--|
| FT  | DTSLPFD  | 106     | 106       | INTERCHAIN (WITH A HEAVY CHAIN).         |
| FT  | DTSLPFD  | 109     | 109       | INTERCHAIN (WITH A HEAVY CHAIN).         |
| FT  | DTSLPFD  | 140     | 200       |  |
| FT  | DTSLPFD  | 246     | 304       |  |
| FT  | SITE     | 156     | 156       | AT OR NEAR THE COMPLEMENT-BINDING SITE.  |
| FT  | MOD_RES  | 326     | 326       | REMOVED POST-TRANSLATIONALLY (PROBABLE). |
| FT  | VARIANT  | 60      | 60        | S -> A (IN MYELOMA PROTEIN TIL & ZIE).   |
| FT  |          |         |           | /FTID=VAR_003889.                        |
| SEQ | SEQUENCE | 326 AA; | 35884 MW; | 8310878C6878CF9C CRC64;                  |

  

|                       |              |                    |                |             |         |
|-----------------------|--------------|--------------------|----------------|-------------|---------|
| Query Match           | 74.7%;       | Score 1158.5;      | DB 1;          | Length 326; |         |
| Best Local Similarity | 68.0%;       | Pred. No. 1.1e-81; |                |             |         |
| Matches 227;          | Conservative | 8;                 | Mismatches 22; | Indels 77;  | Gaps 5; |

  

|    |     |   |     |
|----|-----|---|-----|
| QY | 7   | APCSRGSSWSADLKDCKMDCASCRRPHSDFCGLCAAP--PAPRLLWRS-----       | 54  |
| DB | 12  | APCSRSTS-----ESTALGCLVDYPPPTVTVSNWNGALTSQVHTF               | 53  |
| QY | 55  | -----CDKTH-----TCPPCPAPE                                    | 68  |
| DB | 54  | PAVLQSSGLSYLSLWVTPSSNFGTYTCNVHDKPSNTKVDKTVKRCVCEPCPPAPP     | 113 |
| QY | 69  | ABGAPSVFLPPPKDITLMSRTEPVCVVVDVSHEDPEVKFNWYDGVGVHNAKTKPRE    | 128 |
| DB | 114 | VAG-PSVFLFPKKPKDITLMSRTEPVCVVVDVSHEDPEVQFNWYDGVGVHNAKTKPRE  | 172 |
| QY | 129 | EQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIISKAKGQPREPOVYTLPP | 188 |
| DB | 173 | EQNFSTRVVSVLTVLHQDWLNGKEYCKVSNKGLPAPIEKTIISKTKGQPREPOVYTLPP | 232 |
| QY | 189 | SREMTKNQVSLTCLVKGFFPSDIAVWESNGQPNENYKTPVPVLDSDGSEFFLYSKLTVD | 248 |
| DB | 233 | SREMTKNQVSLTCLVKGFFPSDIAVWESNGQPNENYKTPPMLDSDGSEFFLYSKLTVD  | 292 |
| QY | 249 | KSRWQGNVFSVCSYMHALNNHYTKLSLSPGK                             | 282 |
| DB | 293 | KSRWQGNVFSVCSYMHALNNHYTKLSLSPGK                             | 326 |

  

|           |   |      |         |
|-----------|---|------|---------|
| RESULT 3  |   |      |         |
| GC3_HUMAN | STANDARD;   | PRT; | 290 AA. |
| ID        | GC3_HUMAN   |      |         |
| AC        | P01860;   |      |         |
| DT        | 21-JUL-1986 (Rel. 01, Created)                                      |      |         |
| DT        | 21-JUL-1986 (Rel. 01, Last sequence update)                         |      |         |
| DT        | 20-AUG-2001 (Rel. 40, Last annotation update)                       |      |         |
| DE        | IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).      |      |         |
| GN        | IGHG3.  |      |         |
| OS        | Homo sapiens (Human).   |      |         |
| OC        | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |      |         |
| OC        | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.          |      |         |
| OX        | NCBI_Taxid=9606;  |      |         |
| RNA       | [1]   |      |         |
| SEQ       | SEQUENCE (DISEASE PROTEIN WIS).                                     |      |         |
| RP        | MEDLINE=81021548; PubMed=6774747;                                   |      |         |
| RA        | Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;             |      |         |
| RT        | "Primary structure of human gamma 3 immunoglobulin deletion mutant: |      |         |
| RL        | gamma 3 heavy-chain disease protein Wis.";                          |      |         |
| RL        | Biochemistry 19:4304-4308(1980).                                    |      |         |
| RNA       | [2]   |      |         |
| REV       | REVISIONS TO 12-97 OF PROTEIN WIS.                                  |      |         |
| RP        | MEDLINE=771118561; PubMed=402363;                                   |      |         |
| RA        | Michaelson T.E., Frangione B., Franklin E.C.;                       |      |         |
| RT        | "Primary structure of the 'hinge' region of human IgG3. Probable    |      |         |
| RL        | quadruplication of a 15-amino acid residue basic unit.";            |      |         |
| RL        | J. Biol. Chem. 252:883-889(1977).                                   |      |         |
| RNA       | [3]   |      |         |
| REV       | REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).           |      |         |
| RP        | MEDLINE=77021516; PubMed=823945;                                    |      |         |
| RA        | Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;       |      |         |
| RT        | "The amino acid sequence of 'heavy chain disease' protein ZUC.      |      |         |
| RL        | Structure of the Fc fragment of immunoglobulin G3.";                |      |         |







```

QY 55 -----CDKTH-----TC--PPC 64
Db 46 LTNGVTFPSVROSSGLYSLSSVSVSSOPVTCNVAHPATNTKVDKTVAPSTCSKPTC 105
QY 65 PAPAEGAPSVFLPPPKKDTLMISRTPEVTVVVDVSHEDPEYKFNMYVDGVEVHNNAKT 124
Db 106 PPELLGGPSVFIFFPPPKKDTLMISRTPEVTVVVDVSDQDPEVQFTWYINNEQVTRAP 165
QY 125 KPREEQNSYRVVSVLTVLHQDLNCKEYKCKVSNKALPAPIEKTISKAGQPREQVY 184
Db 166 PLREQQNSYRVVSVLTVLHQDLNCKEYKCKVSNKALPAPIEKTISKARGOPLEPKVY 225
QY 185 TLPPSREEMTKNOVSLTCLVKGYFSPDIAVWESNGOPENNYKTPPVLDSDGSEFLYSK 244
Db 226 TMGPPRELSRSVSLTCMNGFYPSDISEWKEKNGRAEDNYKTPPAVLSDSGSYFLYNK 285
QY 245 LTVDKSRMGOQNFVSCSVMHHEALNNHYTKQSLSLSPGK 282
Db 286 LSVPTSEWQRGDVTCSVMHEALNNHYTKQSLSLSPGK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE IG GAMMA-2 CHAIN C REGION.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-58.
RX MEDLINE=71058471; PubMed=5538606;
RA Birstein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). 3. Amino acid sequence of the region around the
half-cysteine joining heavy and light chains."
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
and hinge region cyanogen bromide fragments."
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
antibodies."
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;
RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
antibodies."
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin."

```

```

RL Biochemistry 10:26-31(1971).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR: A02151; G2GP
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 56.5%; Score 876; DB 1; Length 329;
Best Local Similarity 54.1%; Pred. No. 4.3e-60;
Matches 171; Conservative 28; Mismatches 57; Indels 60; Gaps 4;

QY 26 ASRARPHSDFCLGCAAAP--PAFFRLWRS----- 54
Db 14 ASCDVTSGSMMLGCLVKGYFPEPVTVKNWNSGALTSGVHTFPAVLQSLGSLYSLTSMVTVP 73
QY 55 -----CDKTH-----TCPPCAPEAEGAPSVFLPPPKKDTLM 87
Db 74 SQRATCNVAHPASSTKVDKTVETPTZPBPCCTCPKPPENLGGPSVFIFFPPKPKDTLM 133
QY 88 ISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQNSYRVVSVLTVLHQD 147
Db 134 ISLTPTVTCVVVDVSDQDEPEVQFTWFDNKPVGNAETKPRVEQYNTTFRVESVLPQHOD 193
QY 148 WLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPVITLPPSRDEEMTKNQVSLTCLVKGF 207
Db 194 WLRGKEPKCKVSNKALPAPIEKTISKAKGQPRMPDVITLPPSRDELSKSVSVTCLINF 253
QY 208 YPSDIAVWESNGOP--ENNYKTPPVLDSDGSEFLYSKLTVDKSRMGOQNFVSCSVNHE 265
Db 254 FPADIHVEWASNRVPVSEKEYKNTPPIEDADAGSIFLYSKLTVDKSAWDQGVITCSVNHE 313
QY 266 ALNHHYTKQSLSLSPG 281
Db 314 ALNHHVTQKAISRSPG 329

RESULT 7
GC1_RAT STANDARD; PRT; 326 AA.
ID GC1_RAT
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."
RL Gene 74:473-482(1988).
DR PIR: PS0017;
DR InterPro: IPR003006; Ig_MHC.

```

```
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 54.0%; Score 838; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 3.4e-57;
Matches 150; Conservative 43; Mismatches 47; Indels 8; Gaps 3;

QY 40 CAAPAPPE-----RLLRSCDKTHTCPAPAEAGAPSVFLFPKPKDTLMISRTPEV 94
DB 82 CNVAHPASTKVDKIVPRNCG--GDCKPCICTGSE--VSSVFIFPPKPKDVITLTTPKV 138
QY 95 TCVVVDVSHEDPEVKFNMYDGVGVHNAKTKPREQYNSTYRVSVLVTLVHODWLNKREY 154
DB 139 TCVVVDISQDDPEVHFVDFVDDVEVHTAQTDRPPEQFNSTFRSSELPILHODWLNKRTF 198
QY 155 KCKVSNKALPAPIETISKAKQPREPOVYITLPPSRREMTKNQVSLTCLVKGFFYPSDTAV 214
DB 199 RCKVTSAAFPSPDIETISKPEGRTPQVHYTNMSPTEKEMTQNEVSIITCMVKGFYPPDIYV 258
QY 215 EWESNGQPNKKYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFSVCVYMHAEALHNHYTQK 274
DB 259 ENQMNGQPNQYKNTPTPTMDTDCSFLYSKLVNKKVKEWQGNFTFCVSLHGLHNHHTK 318
QY 275 SLSLSPGK 282
DB 319 SLSHSPGK 326

RESULT 8
GC1_MOUSE
ID GC1_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; J00451; -: NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; IGC1; 2.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 53.9%; Score 836; DB 1; Length 329;
Best Local Similarity 62.7%; Pred. No. 4.9e-57;
Matches 156; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

QY 40 CAAPAPPEPRLLRSCD-----KTHTCP--PCPAPEAEAGAPSVFLFPKPKDTLMISRTPE 93
DB 81 CNVAHPASKTELKRIEPRIPKPTPPGSSCPPGNILGSPSVFIPTPKPKDALMISLTPEK 140
QY 94 VTCVVVDVSHEDPEVKFNMYDGVGVHNAKTKPREQYNSTYRVSVLVTLVHODWLNKRE 153
DB 141 VTCVVVDVSEDDPDVHVSFVNDKVEHTAQTDRPPEAQYNSTFRVVSALPIQHDWNRGKE 200
QY 154 YCKVSNKALPAPIETISKAKQPREPOVYITLPPSRREMTKNQVSLTCLVKGFFYPSDIA 213
DB 201 FKCKVNNKALPAPIETISKPKGKRAQTPOVYITTPPREQMSKKVSLTCLVTFNFSEAS 260
QY 214 VWESNGQPNKKYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFSVCVYMHAEALHNHYTQ 273
DB 261 VWERNGELEDQYKNTPTPTLDSDGYFLYSKLTVDTSWLGQEIFTCVSVVHEALHNHTQ 320
QY 274 KSLSLSPGK 282
DB 321 KNLSRSPGK 329

RESULT 9
GC1_MOUSE
ID GC1_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
RT cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
```

103  
RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
RX MEDLINE=80012837; PubMed=113776;  
RA Rogers J., Clarke P., Salsner W.;  
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
RL heavy chain.";  
RL Nucleic Acids Res. 6:3305-3321(1979).  
[4]  
RN SEQUENCE (MYELOMA PROTEIN MOPC 21).  
RP MEDLINE=78242288; PubMed=98524;  
RA Adetugbo K.;  
RT "Evolution of immunoglobulin subclasses. Primary structure of a  
RL murine myeloma gammal chain.";  
RL J. Biol. Chem. 253:6068-6075(1978).  
[5]  
RN DISULFIDE BONDS (MOPC 21).  
RP MEDLINE=73008889; PubMed=5073237;  
RA Svasti J., Milstein C.;  
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
RL Biochem. J. 126:837-850(1972).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; V00793; CAA24172.1; -;  
DR EMBL; V00793; CAA24173.1; -;  
DR EMBL; V00793; CAA24174.1; -;  
DR EMBL; V00793; CAA24175.1; -;  
DR EMBL; V00793; CAA24176.1; -;  
DR PIR; A02159; GIMS.  
DR GlycosuitedB; P01868; -;  
DR MGD; MGI:96446; Igh-4.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003597; Ig\_cl.  
DR Pfam; PF00047; ig; 3.  
DR SMART; SM00407; Igc1; 2.  
DR PROSITE; PS00290; Ig\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Alternative splicing.  
FT NON\_TER 1 1  
FT DOMAIN 1 97 CH1.  
FT DOMAIN 98 110 HINGE.  
FT DOMAIN 111 217 CH2.  
FT DOMAIN 218 324 CH3.  
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 138 198 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 174 174 /FTID-CAR\_000055.  
FT DISULFID 244 302  
FT MOD\_RES 324 324 REMOVED POST-TRANSLATIONALLY.  
FT CONFLICT 276 276 N -> D (IN REF. 3).  
FT CONFLICT 278 278 N -> D (IN REF. 3).  
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;  
  
Query Match 53.7%; Score 833; DB 1; Length 324;  
Best Local Similarity 55.3%; Pred. No. 8.le-57;  
Matches 151; Conservative 48; Mismatches 60; Indels 14; Gaps 4;  
  
QY 17 ADLKDCKDCASCRRPFDCLGCAAPPAPE-----RLLRSCDKTKHTCP--CPAPEA 69  
Db 59 SDLYTLSSSVTVSPRSTVTNCVAHPASSYKVDKIVPRDGG-----CKPCICTVPEV 114  
QY 70 EGAPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREE 129

115 S---SVFIPPKPKDVLITLTTPKVTCCVVVDISKDDPEQVQSFVDDVEVHTAQTPREE 171  
QY 130 QYNSTRVRSVLTVLHODWLNGLNGEYKCKVSNKALPAPIETISKAKGQPREQVYILPPS 189  
Db 172 QFNSTFRSVELPIMHODWLNGLNGEYKCKVSNKALPAPIETISKAKGQPREQVYILPPS 231  
QY 190 REEMTKNOVSLTCLVKGFPYSDIAVWESNGQPNYKTTTPPVLDSDGSPFLYSLKLTVDK 249  
Db 232 KEQMAKDKVSLTCTMTDFFPEDITVQWNGQPAENKTKTQPMINTNGSIFVYSKLVNQK 291  
QY 250 SRVQGGNVFSCSYNHEALHNHYTKQSLSLSPGK 282  
Db 292 SNWEAGNTFTCSVLHGLHNHTEKSLSHSPGK 324  
  
RESULT 10  
GC1M\_MOUSE STANDARD; PRT; 393 AA.  
AC P01869;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=80045036; PubMed=115593;  
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
RA Takahashi N., Mano Y.;  
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
RL gamma 1 chain gene.";  
RL Cell 18:559-568(1979).  
RN [2]  
RP SEQUENCE OF 323-393 FROM N.A.  
RX MEDLINE=82197626; PubMed=6804950;  
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;  
RT "mRNA for surface immunoglobulin gamma chains encodes a highly  
RT conserved transmembrane sequence and a 28-residue intracellular  
RT domain.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).  
RN [3]  
RP SEQUENCE OF 323-366 FROM N.A.  
RX MEDLINE=82115295; PubMed=6799207;  
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,  
RA Eisenberg D., Wall R.;  
RT "Gene segments encoding transmembrane carboxyl termini of  
RL immunoglobulin gamma chains.";  
RL Cell 26:19-27(1981).  
RN [4]  
RP SEQUENCE OF 1-44 FROM N.A.  
RX MEDLINE=82222190; PubMed=6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
RL immunoglobulin gamma chains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
CC IN SEPARATE EXONS THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
-----
CC EMBL: V00793; CAA24172.1; -
DR EMBL: V00793; CAA24173.1; -
DR EMBL: V00793; CAA24174.1; -
DR PIR: B02159; GIMSI.
DR MGD: MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 53.4%; Score 828; DB 1; Length 393;
Best Local Similarity 55.1%; Pred. No. 2.4e-56;
Matches 150; Conservative 48; Mismatches -60; Indels 14; Gaps 4;

QY 17 ADLRCMDCASCARPHSDFCLGCAAPAPF-----RLLRSCDKHTPCPP--CPAPRA 69
DB 59 SLDYLSSTVTPSSPRSETVTCNVAHPASSTKVDKIVPRDCG---CKPCICTVPEV 114

QY 70 EGAPSVFLPPPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWYGVGVHNAKTPREE 129
DB 115 S---SVFIFFPPKPKDVLITLTPFKVTCVVVDISKDDPEVQFSWFVDVVEVHTAQTPREE 171

QY 130 QYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAIEKTIKSKAGQPREQVYTLPPS 189
DB 172 QFNSTFRSVSELPIMHQDWLNKEPKRVNSAAFPAPIEKTIKSKGRKAPQVYTIPPP 231

QY 190 REEMTKNOVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDK 249
DB 232 KQMAKDKVSLTCLMIDFFPEDITVEWQWNGQPAENYKNTQPIMTNGSYFYVSKLNVQK 291

QY 250 SRWQOGNVFSCVMHEALHNHYTQKSLSLSPG 281
DB 292 SNWEAGNTFTCSVLHGLNHNHTEKSLSHSPG 323

RESULT 11
GCB_RAT ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
CC "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RT
```

```
Gene 74:473-482(1988).
RL PIR: PS0018; PS0018.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR SMART: SM00410; Igcl; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 53.2%; Score 825; DB 1; Length 333;
Best Local Similarity 64.0%; Pred. No. 3.4e-56;
Matches 146; Conservative 33; Mismatches 49; Indels 0; Gaps 0;

QY 55 CDKTHCTCPAPAPAEAGAPSVFLPPPKPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWY 114
DB 106 CPTCTCHKCPVPELIGGSPVFIFPPKPKDILLISQNAKVTCTVVVDVSEEPDVQFSWF 165

QY 115 DGVVEHNAKTKPREQYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAIEKTIKSKA 174
DB 166 NNVEVHTAQTPREQYNSTFRVVSALPIQHODWMSGKEFKCKVNNKALPSPIEKTIKSKP 225

QY 175 KGQPREPQVYTLPPREEMTKNQVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLD 234
DB 226 KGLVRKQVYVNGVPTEQLTEQTSVLTCLTSGFLPNDIGVWTSNGHIEKNYKNTPEVMD 285

QY 235 SGDSFSLYKLVDSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 282
DB 286 SDGSFFMYSKLVNRSWRDSDRAPFVCSVYHGLNHNHVEKSIKSRPPGK 333

RESULT 12
GCB_MOUSE ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84014483; PubMed=6314258;
RA Konaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; J00451; AAB59655.1; -  
DR EMBL; V01526; CAA24767.1; ALT\_SEQ.  
DR PIR; A02155; G3MSM.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003597; Ig\_cl.  
DR InterPro: IPR003600; Ig\_like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGcl; 2.  
DR SMART; SM00410; IG\_like; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Transmembrane; Alternative splicing.  
FT NON\_TER 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 223 CH2.  
FT DOMAIN 224 327 CH3.  
FT TRANSMEM 346 362 POTENTIAL.  
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
FT CONFLICT 333 333 E -> G (IN REF. 2).  
FT CONFLICT 342 342 E -> Q (IN REF. 2).  
FT CONFLICT 388 388 P -> F (IN REF. 2).  
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 53.2%; Score 825; DB 1; Length 398;  
Best Local Similarity 62.3%; Pred. No. 4.2e-56;  
Matches 154; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

QY 40 CAAAPAPFRLLRSCD-----KTHTCP--PCPAEAGAPSVFLFPKPKDTLMISRTPE 93  
DB 81 CNVAHPASKTELKRIEPRIPKPKSTPGSCPGNIGLGPVSFIFPPKPKDALMISLTPK 140  
QY 94 VTCVVVDVSHEDPPVKENWYVDGVEVINAKTPREOYNSTYRVVSVLTVLHQDWLNGKE 153  
DB 141 VTCVVVDVSEDDPDVHVSFVNDKKEHTAQTQPREAQYNSTFRVSALPIQHQQDWMRGE 200  
QY 154 YKCVSNKALPAPIETKISAKGPREPOVYTLPPSREEMTKNOVSLTCLVKGYPSDIA 213  
DB 201 FKCVNNKALPAPIETKISAKGPREPOVYTLPPSREEMTKNOVSLTCLVKGYPSDIA 260  
QY 214 VEWESGQPNKYKTPPVLDSDGSFLLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQ 273  
DB 261 VEWERNGEQDYKNTPPILDSDGTYFLYSLKLTVDTSWLQGEIFTCSVVHEALHNHTQ 320  
QY 274 KSLSLP 280  
DB 321 KNLRSRSP 327

RESULT 13  
GCC\_RAT ID GCC\_RAT STANDARD; PRT; 329 AA.  
AC P20762;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2C CHAIN C REGION.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8816903; PubMed=3127222;  
RA Brueggemann M., Delmastro-Galfrè P., Waldmann H., Calabi F.;  
RT \*Sequence of a rat immunoglobulin gamma 2c heavy chain constant

RT region cDNA: extensive homology to mouse gamma 3.";  
RL Eur. J. Immunol. 18:317-319(1988).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X07189; CAA30169.1; -  
DR PIR; S00847; S00847.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003597; Ig\_cl.  
DR InterPro: IPR003600; Ig\_like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGcl; 2.  
DR SMART; SM00410; IG\_like; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 222 CH2.  
FT DOMAIN 223 329 CH3.  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 249 307  
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 52.7%; Score 817; DB 1; Length 329;  
Best Local Similarity 53.8%; Pred. No. 1.4e-55;  
Matches 155; Conservative 46; Mismatches 49; Indels 38; Gaps 3;

QY 6 TAPCSRGSWS-----ADLDKCMDCASCRRPHSDFLGCAAPAPFRLLRWS 54  
DB 69 TVPSSTWSTQVTCVSAHPATKSNLIKRIEPRPKPRPTDTC-----S 112  
QY 55 CDKTHCTPCCPAPEAGAPSVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVKENWY 114  
DB 113 CD-----DNLGRPSVFIFPPKPKDILMTLTLPKTVVVDVSEEDPVQFSWF 161  
QY 115 DGEVHNNAKTPREOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKA 174  
DB 162 DNVRFVTAQTQPHQEQLNGTFRVYSTLRIHQDWMMSGKEFKCKVNNKDLPSPIETISK 221  
QY 175 KGQPREPOVYTLPPSREEMTKNOVSLTCLVKGYPSDIAVWESNGQPNKYKTPPVLD 234  
DB 222 RGRARTPQVYTIIPPREQMSKNKVSITCMVTSFYPASISVWERNRNGEQLDNTLPVLD 281  
QY 235 SDGSFFLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCK 282  
DB 282 SDESIFLYSLKSLVDTDSWNRGDIYTCVHEALHNHTQKSLSRSPCK 329

RESULT 14  
GCC\_RAT ID GCC\_RAT STANDARD; PRT; 322 AA.  
AC P20760;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 27 82
FT FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 136 196
FT FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 51.9%; Score 804.5; DB 1; Length 322;
Best Local Similarity 47.0%; Pred. No. 1.2e-54;
Matches 157; Conservative 38; Mismatches 62; Indels 77; Gaps 5;

QY 3 APTGAPCSRGSSWADLKDCMDCASCARPHSDFCIGCAAP--PAPFLLWRS----- 54
Db 12 AGTA-----LKSNWVTGLGCLVKGFFPEVTVTWNSGALSSG 49
QY 55 -----CDKTH-----TCPPCPAPE 68
Db 50 VHTFPAVLQSGLYLTSSVTPSTWSSQAVTCNVAHPASSPKVKIKVIPRECPCGCTG 109
QY 69 AGCAPSVLFPKPKDPTLMISRTPEVTCVVDVSHEDVEKENWYVDGEVINATKPRE 128
Db 110 SE-VSSVFIFPPKTKDVLITLTPTKVTGVVDVDSQNDPEVRFSWFIDVVEVHTAQAPE 168
QY 129 EQYNSTYRVSVLTVLHQLDNLGKEYKCKVSNKALPAPIEKTISKAKGPQPPQVYTLPP 188
Db 169 KQNSTLSRSVSELPVHRDLNKGTFCKVNSCAPPIEKSISKEPTGPGQVYTWAP 228
QY 189 SREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPNKYTTPTPVLDSDGSEFFLYSKLTV 248
Db 229 PREEMTQSOVSTICVMKGFYPDPDIYETENKMGQPOENKPTPTMDTDCGYLYSKLVNK 288
QY 249 KSRWQGNVFCVSMVHEALHNHYTKSLSLSPGK 282
Db 289 KETWQGNFTTCVSLVHGLHNHHTKSLSHSPGK 322

RESULT 15
GCAA_MOUSE
ID GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN Nucleic Acids Res. 8:3143-3155(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RN Nucleic Acids Res. 9:1365-1381(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Ollo R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin: amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -.
DR PIR; A02152; G2MSA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 144 204
FT FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 51.5%; Score 799.5; DB 1; Length 330;
Best Local Similarity 49.7%; Pred. No. 3e-54;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:26 ; Search time 54.85 seconds  
(without alignments)  
752.029 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_309

Perfect score: 1551

Sequence: 1 EQAPGTAPCSRGSSWADLD.....MHEALNHHTYKLSLSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL17:\*
  - 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 841   | 54.2        | 463    | 11 Q99LC4 | Q99LC4 mus musculus |
| 2          | 837   | 54.0        | 437    | 11 Q99LA4 | Q99LA4 mus musculus |
| 3          | 804   | 51.8        | 473    | 11 Q99BL4 | Q99BL4 mus musculus |
| 4          | 798.5 | 51.5        | 473    | 11 Q99L25 | Q99L25 mus musculus |
| 5          | 795.5 | 51.3        | 468    | 11 Q99L31 | Q99L31 mus musculus |
| 6          | 367   | 23.7        | 375    | 4 Q9BS21  | Q9BS21 homo sapien  |
| 7          | 367   | 23.7        | 597    | 4 Q9BU10  | Q9BU10 homo sapien  |
| 8          | 367   | 23.7        | 597    | 4 Q9BQ88  | Q9BQ88 homo sapien  |
| 9          | 304.5 | 19.6        | 129    | 4 Q9NP84  | Q9NP84 homo sapien  |
| 10         | 284.5 | 18.3        | 384    | 4 Q9UP60  | Q9UP60 homo sapien  |
| 11         | 279.5 | 18.0        | 500    | 4 Q9BRV0  | Q9BRV0 homo sapien  |
| 12         | 278.5 | 18.0        | 129    | 11 Q9QW3  | Q9QW3 mus musculus  |
| 13         | 278.5 | 18.0        | 129    | 11 Q9CR75 | Q9CR75 mus musculus |
| 14         | 274.5 | 17.7        | 416    | 4 Q9NPP6  | Q9NPP6 mus musculus |
| 15         | 270   | 17.4        | 487    | 11 Q99KA4 | Q99KA4 mus musculus |
| 16         | 269.5 | 17.4        | 684    | 13 Q90544 | Q90544 ginglymosto  |
| 17         | 269   | 17.3        | 426    | 11 Q9DCD9 | Q9DCD9 mus musculus |
| 18         | 255   | 16.4        | 479    | 11 Q99M22 | Q99M22 mus musculus |
| 19         | 255   | 16.4        | 484    | 11 Q99LA6 | Q99LA6 mus musculus |

|    |       |      |     |           |                     |
|----|-------|------|-----|-----------|---------------------|
| 20 | 185   | 11.9 | 268 | 13 Q90524 | Q90524 ginglymosto  |
| 21 | 184.5 | 11.9 | 130 | 11 Q9D8W4 | Q9D8W4 mus musculus |
| 22 | 181   | 11.7 | 235 | 11 Q99M11 | Q99M11 mus musculus |
| 23 | 180   | 11.6 | 509 | 11 Q9QX57 | Q9QX57 mus musculus |
| 24 | 180   | 11.6 | 513 | 11 P97797 | P97797 mus musculus |
| 25 | 176.5 | 11.4 | 259 | 13 Q90530 | Q90530 ginglymosto  |
| 26 | 174.5 | 11.3 | 509 | 11 Q9WTN4 | Q9WTN4 mus musculus |
| 27 | 173   | 11.2 | 509 | 11 Q98907 | Q98907 mus musculus |
| 28 | 171.5 | 11.1 | 257 | 13 Q90536 | Q90536 ginglymosto  |
| 29 | 170   | 11.0 | 237 | 13 Q90545 | Q90545 ginglymosto  |
| 30 | 168   | 10.8 | 105 | 11 Q99JC1 | Q99JC1 mus musculus |
| 31 | 168   | 10.8 | 506 | 6 Q46632  | Q46632 bos taurus   |
| 32 | 167.5 | 10.8 | 252 | 13 Q90568 | Q90568 ginglymosto  |
| 33 | 166.5 | 10.7 | 267 | 13 Q90529 | Q90529 ginglymosto  |
| 34 | 165   | 10.6 | 238 | 7 Q9MXA2  | Q9MXA2 aulonocara   |
| 35 | 165   | 10.6 | 506 | 6 Q46631  | Q46631 bos taurus   |
| 36 | 162   | 10.4 | 261 | 7 Q19363  | Q19363 sus scrofa   |
| 37 | 160   | 10.3 | 261 | 7 Q62868  | Q62868 sus scrofa   |
| 38 | 157.5 | 10.2 | 237 | 7 Q9MX99  | Q9MX99 aulonocara   |
| 39 | 155   | 10.0 | 261 | 7 Q98263  | Q98263 sus scrofa   |
| 40 | 152.5 | 9.8  | 354 | 4 Q9NQK8  | Q9NQK8 homo sapien  |
| 41 | 150   | 9.7  | 208 | 7 Q9MXA0  | Q9MXA0 aulonocara   |
| 42 | 150   | 9.7  | 260 | 7 P79551  | P79551 homo sapien  |
| 43 | 149.5 | 9.6  | 354 | 4 Q9P1W8  | Q9P1W8 homo sapien  |
| 44 | 149   | 9.6  | 238 | 11 Q99M37 | Q99M37 mus musculus |
| 45 | 147.5 | 9.5  | 248 | 7 Q98044  | Q98044 aulonocara   |

ALIGNMENTS

RESULT 1

Q99LC4  
ID Q99LC4; PRELIMINARY; PRT; 463 AA.  
AC Q99LC4;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003435; AAH03435.1; -.  
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 54.2%; Score 841; DB 11; Length 463;  
Best Local Similarity 60.0%; Pred. NO. 9.5e-68;  
Matches 150; Conservative 42; Mismatches 44; Indels 14; Gaps 4;

|    |     |   |     |
|----|-----|---|-----|
| QY | 40  | CAAAPAPP-----RLWRSCDKTHTCPP--CPAPEAGAPSVLEFPKPKOTLMISRTTP     | 92  |
| DB | 221 | CNVAHPASTKVDKIVPRDCG-----CKPCICTVPEVS---SVFIFFPKPKDVLITLTP    | 273 |
| QY | 93  | EYTCVVVDVSHEDPEVKFNWYVDGVEFHNNAKTKPREEQYNSTYRVVSVLTVLDHQLMNGK | 152 |
| DB | 274 | KVTCVVVDISKDDPEVQSFVDDVEVHTAQTQPREQFNSTFRSVELPIHQDLNKGK       | 333 |
| QY | 153 | EYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDI  | 212 |
| DB | 334 | EFKCRVNSAAFPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDI  | 393 |
| QY | 213 | AVENESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHT    | 272 |
| DB | 394 | TVEWQWNGQPAENYKNTQPTIMDTGSGYFYSKLNVQKSNWEAGNTFTCSVLGHEGLNHT   | 453 |
| QY | 273 | OKSLSLSPGK 282  |     |

```
Db 454 EKSLSHSPGK 463
RESULT 2
Q9R1A4 PRELIMINARY; PRT; 437 AA.
ID Q9R1A4
AC Q9R1A4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Nasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF152372; AAD40243.1; -.
DR MGD; MGI:1924014; 1810060009Rik.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG1; 2.
DR SMART; SM00407; IG1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 54.0%; Score 837; DB 11; Length 437;
Best Local Similarity 59.6%; Pred. No. 2e-67;
Matches 149; Conservative 43; Mismatches 44; Indels 14; Gaps 4;
QY 40 CAAPAPPP-----RLWRSCDKTHTCPP--CPAPEAGAPSVFLFPKPKDTLMISRT 92
Db 195 CNVAHPASSTKVDKKIIPVDCG----CKPCICTVPEVS---SVFIPPKPKDVLITLTP 247
QY 93 EYTCVVDVSHEDPEVKFNKYVDGVVHNAKTRPREQYNSTYRVSVLVFLVHODWLNKG 152
Db 248 KVTCCVVVDISKDDPEVQFSFVDDVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNKG 307
QY 153 EYCKCKVSNKALPAPIETISKAKGPREPOVYTLPPSREEMTKNOVSLTCLVKGFPDSI 212
Db 308 EFKRVNSAFAPIETISKTRKPKAPQVYIIPPPKEQAKDKVSLTCTMTDFFPEDI 367
QY 213 AVESNGSQPNKYKTTTPVLDSDGSPFLYSLKTVKDSRWQOQGNVFCSCVMHEALNHYT 272
Db 368 TVEQWNGQPAENYKNTQPTMDTGDGYSFYVSKLVNQSNEAGNTFTCSVHEGLNHH 427
QY 273 OKSLSLSPGK 282
Db 428 EKNLSHSPGK 437

RESULT 3
Q9D8L4 PRELIMINARY; PRT; 473 AA.
ID Q9D8L4
AC Q9D8L4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 1810060009RIK PROTEIN.
GN 1810060009RIK.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Nasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF152372; AAD40243.1; -.
DR MGD; MGI:1924014; 1810060009Rik.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG1; 2.
DR SMART; SM00407; IG1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 51.8%; Score 804; DB 11; Length 473;
Best Local Similarity 58.1%; Pred. No. 2.1e-64;
Matches 147; Conservative 38; Mismatches 66; Indels 2; Gaps 1;
QY 30 ARPHSDPCLGCAAPAPFPRLRLWRSCDKTHTCPPAPEAGAPSVFLFPKPKDTLMIS 89
Db 223 AHPASSTKVDKKIIEPRVP--ITQNPCLPKCEPCAPADLLGGPSVFIFPKIKDVLMS 280
QY 90 RTEVTCVVDVSHEDPEVKFNKYVDGVVHNAKTRPREQYNSTYRVSVLVFLVHODWL 149
Db 281 LSPMTVCVVVDVSEDDPDVQISWVFNVEVHTAQTPHREDYNTSLRVVSALPTIQHODW 340
QY 150 NGREYCKVSNKALPAPIETISKAKGPREPOVYTLPPSREEMTKNOVSLTCLVKGFP 209
Db 341 SGREFKCKVNNRALPSPIETISKTRKPKAPQVYIIPPPAEEMTKKESLTCMTITFLP 400
QY 210 SDIAVWESNGQPNKYKTTTPVLDSDGSPFLYSLKTVKDSRWQOQGNVFCSCVMHEALN 269
Db 401 AEIAVDWTSNGRTEQNYKNTATVLDSDGSPFYMYSKLRVQKSTWERSGLFACSVVHEGLN 460
QY 270 HYTKSLSLSPGK 282
Db 461 HLTKTITSRLGK 473

RESULT 4
Q99L25 PRELIMINARY; PRT; 473 AA.
ID Q99L25
```

```
AC Q99L25;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AAH03888.1; -.
SQ SEQUENCE 473 AA; 52449 MW; BE98898B7986DA155 CRC64;

Query Match 51.5%; Score 798.5; DB 11; Length 473;
Best Local Similarity 45.7%; Pred. No. 6.6e-64;
Matches 164; Conservative 30; Mismatches 80; Indels 85; Gaps 6;

QY 9 CSRGSS-----WSADLDKMDCAS-----CDKTH-----
DB 115 CSRGSIYGYLYFYDYGQCTTIVSSAKTTAPSVYPLAPVCGDGTGGTGGTGLCLVRG 174
QY 45 --PAPFRLLR-----CDKTH-----
DB 175 YPEPVTLTWNSGSLSSGVHTTTPAVLQSDLYTLSSSVTVTSSTWPSQITCNVAHPASST 234
QY 60 -----TCPP--CPAPEAGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSH 103
DB 235 KVDDKIEPRGPTIKCPCKPCAPNLLGGPSVFIIPPKIKDVLMSLSPMTVCVVVDVSE 294
QY 104 EPEVKFNVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKAL 163
DB 295 DDPDQISWFWNNVEVLTQOTHTREDYNSTLRVVSALPIQHDWMSGKEFKCKVNNKAL 354
QY 164 PAPIETKISKAGOPREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 223
DB 355 PAPIETKISKGSVRAPQVYVLPPEEEMTKQVLTCTCMVTFDMPEDIVYVEMTNGKTE 414
QY 224 NNYKTPPVLDSGDFLYSKLTVDKSRWQGNVFSQVHHEALHNHYTKSLSPGK 282
DB 415 LNYKTEPVLDSGDFLYSKLTVDKSRWQGNVFSQVHHEALHNHYTKSLSPGK 282

RESULT 5
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AAH03878.1; -.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B332ADB CRC64;

Query Match 51.3%; Score 795.5; DB 11; Length 468;
Best Local Similarity 49.7%; Pred. No. 1.2e-63;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;

QY 28 CRARPHSDFCLGCAAP--PAPFRLLR-----
DB 153 CGDTTGSSTVTLGCLVKGFYPEPVTLTWNSGSLSSGVHTTTPAVLQSDLYTLSSSVTVTSST 212
```

```
QY 55 -----CDKTH-----TCPP--CPAPEAGAPSVFLPFPKPKDTL 86
DB 213 WFSQSTICNVNVAHPASSTKVDDKIEPRGPTIKCPCKPCAPNLLGGPSVFIIPPKIKDVL 272
QY 87 MISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHQ 146
DB 273 MISLSPMTVCVVVDVSEDDPDQISWFWNNVEVLTQOTHTREDYNSTLRVVSALPIQHQ 332
QY 147 DWLNKGYCKVSNKALPAPIETKISKAGOPREPOVYTLPPSREMTKNQVSLTCLVKG 206
DB 333 DWMSGKEFKCKVNNKALPAPIETKISKGSVRAPQVYVLPPEEEMTKQVLTCTCMVTD 392
QY 207 FYPSDIAVEWESNGOPENNYKTPPVLDSGDFLYSKLTVDKSRWQGNVFSQVHHEA 266
DB 393 FMPEDIYVEMTNGKTELYNKTEPVLDSGDFLYSKLTVDKSRWQGNVFSQVHHEG 452
QY 267 LHNHYTKSLSPGK 282
DB 453 LHNHHTTKSFSRTPGK 468

RESULT 6
Q9BSZ1 PRELIMINARY; PRT; 375 AA.
AC Q9BSZ1;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:10455).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004476; AAH04476.1; -.
SQ SEQUENCE 375 AA; 41314 MW; BLA0A098F473619 CRC64;

Query Match 23.7%; Score 367; DB 4; Length 375;
Best Local Similarity 33.8%; Pred. No. 3.8e-25;
Matches 75; Conservative 50; Mismatches 89; Indels 8; Gaps 6;

QY 67 PEAGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTKP 126
DB 138 PQDQTAIRVFAIPPS-FASIFLTKSTKLCVLTVDLTLYD-SVTISWTRONGEAVKTHNI 195
QY 127 REQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIETKISKAGOP-REPOVYT 185
DB 196 SESHNPATFSAVGEASICEDDWNGBRFTCTVTHTDLPSPKQITSRPKGVALHRPDVYL 255
QY 186 LPPSREMT-KNOVSLTCLVKGFYPSDIAVEWESNGOP--ENNYKTPPVLD--SDGSFF 240
DB 256 LPPAREQLNLRRESATITCLVTGFSPADYFQWNGRQGLSPKPEKYIVTSAPMPEQAPGRYF 315
QY 241 LYSKLTVDKSRWQGNVFSQVHHEALHNHYTKSLSPGK 282
DB 316 AHSILVSEEWNTGETYTCVVAHEALPNRVTERTVTDKSTGK 357

RESULT 7
Q9BU10 PRELIMINARY; PRT; 597 AA.
AC Q9BU10;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:1652).
OS Homo sapiens (Human).
```



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN NCBI\_TaxID=9606;  
 [1]  
 RA SEQUENCE FROM N.A.  
 RP Zheng S., Cao J., Cao W., Cai X., Geng L.;  
 RT "Identification and characterization of SNCT73, a gene which is down-  
 regulated in colorectal cancer."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX  
 CC DOMAIN.  
 DR EMBL: AF067420; AAC19365.1; -.  
 DR HSSP: P01825; 7FAB.  
 DR InterPro: IPR003597; Ig\_c1.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig\_3.  
 DR SMART: SM00407; IgC1; 2.  
 DR PROSITE: PS00290; IG\_MHC; UNKNOWN\_1.  
 SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;  
 Query Match 18.3%; Score 284.5; DB 4; Length 384;  
 Best Local Similarity 26.0%; Pred. No. 1.1e-17;  
 Matches 89; Conservative 38; Mismatches 138; Indels 77; Gaps 12;  
 QY 13 SWSADLDKCMDCASCARPHSDFCILGCAAA---PPAPFLLWRS----- 54  
 DB 30 SSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQELSVTWSESGQGVTAARNFPPSD 89  
 QY 55 -----CDKTHTC-----PPCPAPEAGCAPSVFLFPPK 81  
 DB 90 ASGDLTYTSSQLTLPATQCLACKSVTCHVKHYTNPSQDVTVPCTPSTPTSPST--PT 148  
 QY 82 PK-----DTLMTSRTEVTCVVVDVSHEDPEVKFNKALPAIEKTIKAKQPREPOVYTL 126  
 DB 149 PSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGL-RDASGVTFWTTPSSGK--SAVQGP 205  
 QY 127 REEQNSTYRVYVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTL 186  
 DB 206 PERDLGCGYSVSVLPGCAPENNHGKFTCTAAYPESKTPLTATLSKS-GNTRFPEVHLL 264  
 QY 187 PPSREEMTKNQ-VSLTCLVKGYFSPDSIAVEWESNGO--PENNYKTTTPPVLD-SDG--SFF 240  
 DB 265 PPPSEELALNELVTLTCLARGSPKDVLRWLGSGQLPREKYLITWASRQEPSQGTITFA 324  
 QY 241 LYSKLTVDKSRWQQGVNFCVSMHEALHNHYTQKSLSLSPGK 282  
 DB 325 VTSILRVAEDWKKGDTFCMVGHEALPLAFTQKTDRLAGK 365  
 RESULT 11  
 Q9BRV0  
 ID Q9BRV0 PRELIMINARY; PRT; 500 AA.  
 AC Q9BRV0;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DE UNKNOWN (PROTEIN FOR MGC:14586).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PROSTATE;  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC005951; AAH05951.1; -.  
 SQ SEQUENCE 500 AA; 54154 MW; 0A9BF43F2A3CC6D9 CRC64;  
 Query Match 18.0%; Score 279.5; DB 4; Length 500;  
 Best Local Similarity 25.4%; Pred. No. 4.1e-17;

Matches 87; Conservative 41; Mismatches 137; Indels 77; Gaps 12;  
 QY 13 SWSADLDKCMDCASCARPHSDFCILGCAAA---PPAPFLLWRS----- 54  
 DB 146 SSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQELSVTWSESGQGVTAARNFPPSD 205  
 QY 55 -----CDKTHTC-----PPCPAPEAGCAPSVFLFPPK 81  
 DB 206 ASGDLTYTSSQLTLPATQCLACKSVTCHVKHYTNPSQDVTVPCTPSTPTSPST--PT 264  
 QY 82 PK-----DTLMTSRTEVTCVVVDVSHEDPEVKFNKALPAIEKTIKAKQPREPOVYTL 126  
 DB 265 PSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGL-RDASGVTFWTTPSSGK--SAVQGP 321  
 QY 127 REEQNSTYRVYVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTL 186  
 DB 322 PERDLGCGYSVSVLPGCAPENNHGKFTCTAAYPESKTPLTATLSKS-GNTRFPEVHLL 380  
 QY 187 PPSREEMTKNQ-VSLTCLVKGYFSPDSIAVEWESNGO--PENNYKTTTPPVLD-SDG--SFF 240  
 DB 381 PPPSEELALNELVTLTCLARGSPKDVLRWLGSGQLPREKYLITWASRQEPSQGTITFA 440  
 QY 241 LYSKLTVDKSRWQQGVNFCVSMHEALHNHYTQKSLSLSPGK 282  
 DB 441 VTSILRVAEDWKKGDTFCMVGHEALPLAFTQKTDRLAGK 482  
 RESULT 12  
 Q9QZW3  
 ID Q9QZW3 PRELIMINARY; PRT; 129 AA.  
 AC Q9QZW3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE TYPE I TRANSMEMBRANE PROTEIN FN14.  
 GN FGFRP2 OR FN14.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=20020297; PubMed=10551889;  
 RA Meighan-Mantha R.L., Hsu D.K.W., Guo Y., Brown S.A.N., Feng S.-L.Y.,  
 RA Peifley K.A., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A.,  
 RA Richards C.M., Winkles J.A.;  
 RT "The mitogen-inducible Fn14 gene encodes a type I transmembrane  
 protein that modulates fibroblast adhesion and migration."  
 RL J. Biol. Chem. 274:33166-33176(1999).  
 DR EMBL: AF156164; AAF0782.1; -.  
 DR MGD; MGI:1351484; Fgfrp2.  
 KW Transmembrane.  
 SQ SEQUENCE 129 AA; 13637 MW; 14B5C68EEF493385 CRC64;  
 Query Match 18.0%; Score 278.5; DB 11; Length 129;  
 Best Local Similarity 56.0%; Pred. No. 1.1e-17;  
 Matches 56; Conservative 2; Mismatches 19; Indels 23; Gaps 2;  
 QY 1 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCILGCAAAAPPAPFRL----- 51  
 DB 28 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCILGCAAAAPPAPFRLWLPILGALS 87  
 QY 52 -----WRSCDKTHT--CPFCPAPEAGCAPSVFL 77  
 DB 88 LVLVLAVSSFLWRRRCRRREKFTTTEETGEGCGVAL 127  
 RESULT 13  
 Q9CR75  
 ID Q9CR75 PRELIMINARY; PRT; 129 AA.  
 AC Q9CR75;

```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FBROBLAST GROWTH FACTOR REGULATED PROTEIN 2.
GN FGRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA, AND CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Ono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005530; BAB24101.1; -
DR EMBL; AK005382; BAB23989.1; -
DR MGD; MGI:1351484; Fgfrp2.
SQ SEQUENCE 129 AA; 13641 MW; 1665C68B4D9A253 CRC64;

Query Match 18.08; Score 278.5; DB 11; Length 129;
Best Local Similarity 56.08; Pred. No. 1.1e-17;
Matches 56; Conservative 2; Mismatches 19; Indels 23; Gaps 2;

Qy 1 EQAPGTAPCSRGSSWSADLDCMDCASCARPHSDFCILGCAAAAPPAPFRLI----- 51
Dy 28 EQAPGTSPCSSGSSWSADLDCMDCASCARPHSDFCILGCAAAAPPAPFRLIWPILGALS 87
Qy 52 -----WRSCDKTHT-CPPCPAPAEAGAPSVFL 77
Dy 88 LVLVLALVSSPLVWRRCRRRREKFTTPIETGGEGCPGVAL 127

RESULT 14
Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundberg J.;
RT "The European IMAGE consortium for integrated molecular analysis of
```

```
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AL389978; CAB97534.1; -
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; Ig_like; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 17.78; Score 274.5; DB 4; Length 416;
Best Local Similarity 29.68; Pred. No. 9.3e-17;
Matches 81; Conservative 38; Mismatches 116; Indels 39; Gaps 12;

Qy 21 KCMD-CASCRAR----PHSDFCLGCAAAAPPAPFRLLRSCDKTHTCPPCPAPAEAGAPS 74
Dy 152 QCPGKSVTCHVKHYTNPSQDVTVCPPPPP-----PPCCHPR----- 189
Qy 75 VFLFPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTRPEQYNST 134
Dy 190 LSLHRPALED-LLLGSEANLTCTLTGL-RDASGATFTWTPSSGK--SAVQGPDPDLGCG 245
Qy 135 YRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRERM 194
Dy 246 YSVSVLPGCCAQPNHGETFTCTAAHELPKLTPLTANITKS-GNTPRPEVHLLPPSEELA 304
Qy 195 KNO-VSLTCLVKGYFSPDAIVAEWESNGO--PENNYKTTTPPVLD-SDG--SFFLYSKLTVD 248
Dy 305 LNELVTLCLARGSPKDVLRWLGSSGSELPREKLTWASRQEPSQGTTFATVAILRVA 364
Qy 249 KSRMQQGNVFCVSMVHEALHNYHTQKLSLSPGK 282
Dy 365 AEDMKKGDTFCMVGHEALPLAFTQKTDRLAGK 398

RESULT 15
Q99KA4 PRELIMINARY; PRT; 487 AA.
AC Q99KA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:6727).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004786; AAH04786.1; -
SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

Query Match 17.48; Score 270; DB 11; Length 487;
Best Local Similarity 30.68; Pred. No. 2.8e-16;
Matches 77; Conservative 41; Mismatches 100; Indels 34; Gaps 11;

Qy 40 CAAAPPAPFRLLRSCDKTHTCPPCPAPAEAGAPSVFLFPKPKDTLMISRTPEVTCVV 99
Dy 243 CSGPPP-----PCPPCP-PSCH--PSLSLQRPALD-LLLGSDASLTCTLN 284
Qy 100 DVSHEDPEVKFNKYVDGVEVHNNAKTRPEQYNST---YRVSVLTVLHQDWLNGKEYK 156
Dy 285 GLRNPAGAV-FTW-----EPSTGKDAVQKAVQNSCGYSSVSLPGCAERNWSGAFK 338
```



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:13 ; Search time 53.92 Seconds  
(without alignments)  
71.436 Million cell updates/sec

Title: US-09-742-454a-7\_COPY\_28\_79  
Perfect score: 301  
Sequence: 1 EQAGTAPCSRGSNSADLD.....HSDFCIGCAAPAPFRLW 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101:\*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID          | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 301   | 100.0       | 114    | 20 AAW73409 | Human secreted pro |
| 2          | 301   | 100.0       | 129    | 20 AAW88506 | Human liver clone  |
| 3          | 301   | 100.0       | 129    | 21 AAU57940 | Human transmembran |
| 4          | 301   | 100.0       | 129    | 22 AAU03498 | Human TWEAK recept |
| 5          | 301   | 100.0       | 309    | 22 AAU03500 | Human TWEAK recept |
| 6          | 147.5 | 49.0        | 112    | 21 AAY91463 | Human secreted pro |
| 7          | 147.5 | 49.0        | 155    | 21 AAY91604 | Human secreted pro |
| 8          | 147.5 | 49.0        | 156    | 21 AAY91552 | Human secreted pro |
| 9          | 67.5  | 22.4        | 242    | 22 AAG92250 | C glutamicum prote |
| 10         | 67    | 22.3        | 928    | 17 AAR97853 | Rat REK7 eph-relat |
| 11         | 67    | 22.3        | 1005   | 20 AAW83147 | Rat receptor tyros |

|    |      |      |      |    |          |                    |
|----|------|------|------|----|----------|--------------------|
| 12 | 65.5 | 21.8 | 187  | 20 | AAV29189 | Amino acid sequenc |
| 13 | 64   | 21.3 | 77   | 20 | AAV12612 | Human 5' EST secre |
| 14 | 64   | 21.3 | 115  | 20 | AAV59693 | Secreted protein 4 |
| 15 | 64   | 21.3 | 115  | 21 | AAB08524 | Protein encoded by |
| 16 | 64   | 21.3 | 115  | 21 | AAV65397 | Human 5' EST relat |
| 17 | 64   | 21.3 | 115  | 22 | AAG93299 | Human protein HP10 |
| 18 | 64   | 21.3 | 120  | 21 | AAV53361 | Human colon cancer |
| 19 | 64   | 21.3 | 141  | 20 | AAV73883 | Human prostate tum |
| 20 | 64   | 21.3 | 141  | 20 | AAV59770 | Human normal ovari |
| 21 | 64   | 21.3 | 478  | 22 | AAB48033 | Egf receptor relat |
| 22 | 62.5 | 20.8 | 405  | 19 | AAW33737 | Epidermal growth f |
| 23 | 62.5 | 20.8 | 509  | 14 | AAR38210 | LD2D3D4 EGF recept |
| 24 | 62.5 | 20.8 | 529  | 14 | AAR38209 | LD1D2D3.ApAl EGF r |
| 25 | 62.5 | 20.8 | 621  | 21 | AAV67309 | Epidermal growth f |
| 26 | 62.5 | 20.8 | 1210 | 21 | AAV19259 | Amino acid sequenc |
| 27 | 62.5 | 20.8 | 1210 | 21 | AAV50616 | Human EGF receptor |
| 28 | 62.5 | 20.8 | 1210 | 22 | AAB68420 | Amino acid sequenc |
| 29 | 62.5 | 20.8 | 1251 | 16 | AAR79475 | Mouse LTBP-3. Mus  |
| 30 | 62.5 | 20.8 | 1251 | 22 | AAB61481 | Murine LTBP-3 prot |
| 31 | 62.5 | 20.8 | 1251 | 22 | AAB61483 | Human TANGO 300 ex |
| 32 | 62.5 | 20.8 | 3075 | 19 | AAW50892 | Human laminin A ch |
| 33 | 62   | 20.6 | 234  | 22 | AAW24485 | Human EST encoded  |
| 34 | 62   | 20.6 | 567  | 22 | AAV70531 | Human PRO1 protein |
| 35 | 62   | 20.6 | 720  | 21 | AAV88280 | Human TANGO 215 pr |
| 36 | 62   | 20.6 | 720  | 21 | AAV66695 | Membrane-bound pro |
| 37 | 62   | 20.6 | 720  | 22 | AAU00401 | Human secreted pro |
| 38 | 62   | 20.6 | 720  | 22 | AAV87544 | Human PRO1344. Ho  |
| 39 | 62   | 20.6 | 720  | 22 | AAV70532 | Human PRO2 protein |
| 40 | 62   | 20.6 | 720  | 22 | AAB65218 | Human PRO1344 (UNQ |
| 41 | 62   | 20.6 | 737  | 22 | AAV93670 | Human protein sequ |
| 42 | 61   | 20.3 | 392  | 21 | AAV38394 | Human secreted pro |
| 43 | 60.5 | 20.1 | 271  | 21 | AAV19706 | Protease-resistant |
| 44 | 60.5 | 20.1 | 271  | 21 | AAV19707 | Protease-resistant |
| 45 | 60.5 | 20.1 | 271  | 22 | AAV68046 | Amino acid sequenc |

ALIGNMENTS

|          |  |
|----------|--|
| RESULT   | 1  |
| AAW73409 |  |
| ID       | AAW73409 standard; Protein; 114 AA.  |
| AC       | AAW73409;  |
| XX       |  |
| DT       | 19-FEB-1999 (first entry)  |
| XX       |  |
| DE       | Human secreted protein encoded by Gene No. 13.                             |
| XX       |  |
| KW       | Secreted protein; human; protein therapy; gene therapy; blood disorder;    |
| KW       | pathological condition; diagnosis; cancer; neurological disorder; disease; |
| KW       | developmental abnormality; foetal deficiency; leukaemia; hepatic disease;  |
| KW       | immune system disorder; Alzheimer's disease; cognitive disorder;           |
| KW       | schizophrenia; prostate disease; autoimmune disorder; AIDS.                |
| OS       | Homo sapiens.  |
| XX       |  |
| FH       | Key  |
| FT       | Misc-difference 114  |
| FT       | /note= "unspecified amino acid"  |
| XX       |  |
| PN       | WO9854206-A1.  |
| XX       |  |
| PD       | 03-DEC-1998.   |
| XX       |  |
| PF       | 28-MAY-1998; 98WO-US10868.   |
| XX       |  |
| PR       | 29-AUG-1997; 97US-0056296.   |
| PR       | 30-MAY-1997; 97US-0044039.   |
| PR       | 30-MAY-1997; 97US-0048093.   |
| PR       | 30-MAY-1997; 97US-0048101.   |
| PR       | 30-MAY-1997; 97US-0048190.   |
| PR       | 30-MAY-1997; 97US-0048356.   |

PR 30-MAY-1997; 97US-0050935.  
PR 29-AUG-1997; 97US-0056250.  
PR 29-AUG-1997; 97US-0056293.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Carter KC, Dillon PJ, Endress GA, Feng P, Ni J;  
PI Rosen CA, Ruben SM, Yu G;  
XX WPI; 1999-070209/06.  
DR N-PSDB; AAV08823.  
XX  
XX New isolated human genes - useful for diagnosis and treatment of,  
PT e.g. cancers, neurological disorders, immune diseases, developmental  
PT disorders or blood disorders  
XX  
XX Claim 11; Page 153; 188pp; English.  
XX  
XX This sequence is encoded by a cDNA of the invention, designated  
CC Gene No.13. This sequence represents a human secreted protein, and is  
CC expressed in keratinocytes and to a lesser extent in endothelial  
CC cells and placenta.  
CC The DNA sequences of the invention and their corresponding secreted  
CC polypeptides are useful for preventing, treating or ameliorating medical  
CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
CC can be diagnosed by determining the amount of the new polypeptides in a  
CC sample or by determining the presence of mutations in the DNA sequences.  
CC Specific uses are described for each of the DNA sequences and the encoded  
CC proteins, based on which tissues they are most highly expressed in, and  
CC include developing products for the diagnosis or treatment of cancer,  
CC tumours, neurological disorders, developmental abnormalities and foetal  
CC deficiencies, blood disorders, leukaemias, diseases of the immune system  
CC (including allergies or asthma), hepatic disease, Alzheimer's and  
CC cognitive disorders, schizophrenia, prostate diseases, autoimmune  
CC disorders and AIDS. The polypeptides are also useful for identifying  
CC their binding partners.  
XX  
XX Sequence 114 AA;  
SQ  
Query Match 100.0%; Score 301; DB 20; Length 114;  
Best Local Similarity 100.0%; Pred. No. 3e-26;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGCAAAPAPFRLLW 52  
Db 28 eqapgtapcsrgsswsadldkcmdcascrarphsd fclgcaaaappapfrllw 79  
RESULT 2  
AAW88506  
ID AAW88506 standard; Protein; 129 AA.  
XX  
XX AAW88506;  
XX  
XX 30-MAR-1999 (first entry)  
XX  
XX Human liver clone HP10432-encoded membrane protein.  
XX  
XX Transmembrane protein; HP10432; human; liver.  
XX  
XX Homo sapiens.  
XX  
XX WO9855508-A2.  
XX  
XX 10-DEC-1998.  
XX  
XX 03-JUN-1998; 98WO-JP02445.  
XX  
XX 03-JUN-1997; 97JP-0144948.  
XX  
XX (PROT-) PROTEGENE INC.  
PA (SAGA) SAGAMI CHEM RES CENTRE.

XX Kato S, Sekine S, Yamaguchi T;  
PI WPI; 1999-045730/04.  
XX N-PSDB; AAV84374.  
XX  
XX New human proteins containing transmembrane domains and their  
PT encoding sequences - useful in the preparation of antibodies and  
PT large-scale protein production, gene diagnosis, and gene therapy  
XX  
XX Claim 1; Page 152-153; 178pp; English.  
XX  
XX This is the amino acid sequence of a transmembrane protein encoded  
CC by human liver cDNA clone HP10432 (see AAV84374). The encoded protein  
CC has a signal-like N-terminal region and one internal transmembrane  
CC domain. The invention provides nucleotide sequences (see AAV84359-76)  
CC coding for 18 transmembrane proteins (see AAV8491-508), vectors  
CC containing such polynucleotides, and eukaryotic cells containing the  
CC vectors. The proteins can be used as antigens or as compositions  
CC in the preparation of antibodies against the proteins. The  
CC polynucleotides can be used as probes for gene diagnosis, and as  
CC gene sources for gene therapy and large-scale production of proteins  
CC encoded by the cDNA. The host cells are used for the detection of  
CC ligands corresponding to the expressed proteins, and the screening  
CC of low mol.wt. medicines.  
XX  
XX Sequence 129 AA;  
SQ  
Query Match 100.0%; Score 301; DB 20; Length 129;  
Best Local Similarity 100.0%; Pred. No. 3.4e-26;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGCAAAPAPFRLLW 52  
Db 28 eqapgtapcsrgsswsadldkcmdcascrarphsd fclgcaaaappapfrllw 79  
RESULT 3  
AAW57940  
ID AAY57940 standard; Protein; 129 AA.  
XX  
XX AAY57940;  
XX  
XX 23-MAR-2000 (first entry)  
XX  
XX Human transmembrane protein HTMPN-64.  
XX  
XX Human; transmembrane protein; HTMPN; diagnosis; immunospecific;  
XX antiproliferative; neuroprotective; immune disorder;  
XX reproductive disorder; smooth muscle disorder; neurological disorder;  
XX gastrointestinal disorder; developmental disorder;  
XX cell proliferative disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO9961471-A2.  
XX  
XX 02-DEC-1999.  
XX  
XX 28-MAY-1999; 99WO-US11904.  
XX  
XX 29-MAY-1998; 98US-0087260.  
XX  
XX 02-JUL-1998; 98US-0091674.  
XX  
XX 02-OCT-1998; 98US-0102954.  
XX  
XX 24-NOV-1998; 98US-0109869.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;  
PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;  
PI Au-Young J;  
XX

DR WPI: 2000-072605/06.  
DR N-PSDB: AAZ56761.  
XX  
XX Proteins, polynucleotides, vectors, host cells and antibodies used to  
PT diagnose, treat or prevent immune, reproductive, smooth muscle,  
PT neurological, gastrointestinal, developmental and cell proliferative  
PT disorders -  
XX  
XX  
PS Claim 1; Page 163; 229pp; English.  
XX  
XX AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human  
CC transmembrane proteins designated HTPMN-1 to HTPMN-79, respectively.  
CC The transmembrane protein have immunospecific, antiproliferative and  
CC neuroprotective activities. The human transmembrane proteins,  
CC polynucleotides encoding them and other compositions and methods from  
CC the present invention, can be used for the diagnosis, treatment or  
CC prevention of immune, reproductive, smooth muscle, neurological,  
CC gastrointestinal, developmental and cell proliferative disorders. The  
CC HTPMN's can be used to treat or prevent disorders associated with a  
CC decreased expression or activity of HTPMN.  
XX  
XX Sequence 129 AA;  
SQ  
  
Query Match 100.0%; Score 301; DB 21; Length 129;  
Best Local Similarity 100.0%; Pred. No. 3.4e-26;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 EQAGTAPCSRGSSWSADLCKMDCASCRRPHSDFCLGCAAAAPPAPFRLW 52  
Db 28 eqagtapcsrgsswsadldckmcdascrrphsdfclgcaaaappapfllw 79  
  
RESULT 4  
AAU03498  
ID AAU03498 standard; Protein: 129 AA.  
XX  
XX AAU03498;  
XX  
XX  
DT 26-SEP-2001 (first entry)  
XX  
XX  
DE Human TWEAK receptor (TWEAKR) polypeptide.  
XX  
XX TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;  
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
KW corneal graft neovascularisation; psoriasis; metastatic condition;  
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
KW peripheral atherosclerosis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Location/Qualifiers  
FH Key 1..27  
FT Peptide /note= "Signal peptide"  
FT Protein 28..129  
FT FT /note= "Mature human TWEAKR protein"  
FT Domain 1..78  
FT FT /note= "Extracellular domain"  
FT Domain 79..101  
FT FT /note= "Transmembrane domain"  
FT Domain 102..129  
FT FT /note= "Intracellular domain"  
XX  
XX WO200145730-A2.  
PN  
XX  
XX 28-JUN-2001.  
PD  
XX  
XX 19-DEC-2000; 2000WO-US34755.  
PF

XX 20-DEC-1999; 99US-0172878.  
PR 10-MAY-2000; 2000US-0203347.  
XX  
XX (IMV ) IMMUNEX CORP.  
PA  
XX  
PI Willey SR;  
XX  
XX WPI: 2001-417975/44.  
DR N-PSDB: AAS03963.  
XX  
XX Modulating angiogenesis in a mammal for treating diseases mediated by  
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
PT peripheral tissue, by administering antagonist or agonist of TWEAK  
PT receptor -  
XX  
XX Example 1; Fig 1; 46pp; English.  
XX  
XX The sequence represents the human TWEAK receptor (TWEAKR) protein. The  
CC TWEAK protein is a member of the tumour necrosis factor (TNF) family and  
CC induces angiogenesis. TWEAKR may therefore be used to screen for and  
CC develop TWEAKR agonists and antagonists for the modulation of  
CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
CC The disorders mediated by angiogenesis include ocular disorders  
CC characterised by ocular neovascularisation such as diabetic retinopathy,  
CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration and  
CC corneal graft neovascularisation, and inflammatory diseases such as  
CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
CC malignant and metastatic conditions such as sarcomas and carcinomas,  
CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
XX  
XX Sequence 129 AA;  
SQ  
  
Query Match 100.0%; Score 301; DB 22; Length 129;  
Best Local Similarity 100.0%; Pred. No. 3.4e-26;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 EQAGTAPCSRGSSWSADLCKMDCASCRRPHSDFCLGCAAAAPPAPFRLW 52  
Db 28 eqagtapcsrgsswsadldckmcdascrrphsdfclgcaaaappapfllw 79  
  
RESULT 5  
AAU03500  
ID AAU03500 standard; Protein: 309 AA.  
XX  
XX AAU03500;  
XX  
XX 26-SEP-2001 (first entry)  
XX  
XX Human TWEAK receptor-Fc (TWEAKR-Fc) fusion polypeptide.  
XX  
XX TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;  
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
KW corneal graft neovascularisation; psoriasis; metastatic condition;  
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
KW peripheral atherosclerosis; human IgG1; TWEAKR-Fc; fusion protein.  
XX  
XX Homo sapiens.  
OS  
XX Synthetic.  
XX  
XX Key 1..27  
FH Peptide  
FT Location/Qualifiers



QY 1 EQAPGTAPCGSSWSADLDKMDCA-SC 28  
CC AAY91604  
XX |||||||||||||||||||f: ||  
Db 28 eqapgtapcrgsswsadldkcmdcstsc 56

RESULT 7  
AAY91604  
ID AAY91604 standard; Protein; 155 AA.  
XX  
AC AAY91604;  
XX  
DT 29-JUN-2000 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:277.  
XX  
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
KW antiHIV; antiinflammatory; neutropic; neuroprotective; antiallergic;  
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;  
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;  
KW immune disease; inflammation; blood disorder; tumour; chromosome 16.  
XX  
OS Homo sapiens.  
XX  
PN WO200006698-A1.  
XX  
PD 10-FEB-2000.  
XX  
PF 29-JUL-1999; 99WO-US17130.  
XX  
PR 30-JUL-1998; 98US-0094657.  
PR 05-AUG-1998; 98US-0095486.  
PR 06-AUG-1998; 98US-0095454.  
PR 06-AUG-1998; 98US-0095455.  
PR 12-AUG-1998; 98US-0096319.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;  
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;  
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;  
XX  
WPI; 2000-195282/17.  
XX  
New isolated human genes and the secreted polypeptides they encode,  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders -  
XX  
Disclosure: Page 36-37; 634pp; English.  
XX  
The polynucleotide sequences given in AAA26346 to AAA26458 encode the  
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted  
CC proteins can have activities based on the tissues and cells they are  
CC expressed in. Examples of the activities are: cytostatic;  
CC immunosuppressive; antiHIV; antiinflammatory; neutropic; neuroprotective;  
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;  
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their  
CC corresponding secreted proteins are useful for preventing, treating or  
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also  
CC pathological conditions can be diagnosed by determining the amount of the  
CC proteins in a sample or by determining the presence of mutations in the  
CC polynucleotides, based on which tissues they are most highly expressed  
CC in, and include developing products for the diagnosis or treatment of  
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities  
CC and foetal deficiencies, blood disorders, diseases of the immune system,  
CC autoimmune diseases, hepatic and renal disease, inflammation,  
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,  
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
CC cardiovascular disorders, reproductive disorders, gastrointestinal  
CC disorders, respiratory disorders and metabolic disorders. The proteins  
CC or polynucleotides can also be used as food additives or preservatives.  
CC The proteins are also useful for identifying their binding partners.

CC AAA26337 to AAA26345 and AAY91450 are sequences used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 155 AA;  
Query Match 49.0%; Score 147.5; DB 21; Length 155;  
Best Local Similarity 93.1%; Pred. No. 4.2e-09;  
Matches 27; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
QY 1 EQAPGTAPCGSSWSADLDKMDCA-SC 28  
CC |||||||||||||||||||f: ||  
Db 28 eqapgtapcrgsswsadldkcmdcstsc 56  
RESULT 8  
AAY91552  
ID AAY91552 standard; Protein; 156 AA.  
XX  
AC AAY91552;  
XX  
DT 29-JUN-2000 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:225.  
XX  
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
KW antiHIV; antiinflammatory; neutropic; neuroprotective; antiallergic;  
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;  
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;  
KW immune disease; inflammation; blood disorder; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200006698-A1.  
XX  
PD 10-FEB-2000.  
XX  
PF 29-JUL-1999; 99WO-US17130.  
XX  
PR 30-JUL-1998; 98US-0094657.  
PR 05-AUG-1998; 98US-0095486.  
PR 06-AUG-1998; 98US-0095454.  
PR 06-AUG-1998; 98US-0095455.  
PR 12-AUG-1998; 98US-0096319.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;  
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;  
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;  
XX  
WPI; 2000-195282/17.  
XX  
N-PSDB; AAA26447.  
XX  
New isolated human genes and the secreted polypeptides they encode,  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders -  
XX  
Claim 11; Page 528; 634pp; English.  
XX  
The polynucleotide sequences given in AAA26346 to AAA26458 encode the  
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted  
CC proteins can have activities based on the tissues and cells they are  
CC expressed in. Examples of the activities are: cytostatic;  
CC immunosuppressive; antiHIV; antiinflammatory; neutropic; neuroprotective;  
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;  
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their  
CC corresponding secreted proteins are useful for preventing, treating or  
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also  
CC pathological conditions can be diagnosed by determining the amount of the  
CC proteins in a sample or by determining the presence of mutations in the  
CC polynucleotides. Specific uses are described for each of the  
CC polynucleotides, based on which tissues they are most highly expressed



```
CC (AAT18897).
XX
SQ Sequence 928 AA;

Query Match 22.3%; Score 67; DB 17; Length 928;
Best Local Similarity 31.6%; Pred. No. 20;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDCM-----DCASCR-----ARPHSDFCLGCAAAP 45
Db 273 ppkmhcsaegewlvpgkcmckagyeekngtcqvcprpffkasphsqtcskc---pp 326

RESULT 11
AAW83147
ID AAW83147 standard; Protein; 1005 AA.
XX
AC AAW83147;
XX
DT 11-FEB-1999 (first entry)
XX
DE Rat receptor tyrosine kinase Etk-1.
XX
KW Receptor tyrosine kinase; Ror-1; Ror-2; Etk-1; Etk-2; detection;
KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW binding protein; BDNF; NT-3; diagnosis.
XX
OS Rattus sp.
XX
PN US5843749-A.
XX
PD 01-DEC-1998.
XX
PF 06-JUN-1995; 95US-0469537.
XX
PR 17-MAR-1995; 95US-0406247.
PR 26-JUL-1991; 91US-0736559.
PR 28-OCT-1993; 93US-0144992.
PR 06-JUN-1995; 95US-0469537.
XX
PA (REGG-) REGENERON PHARM INC.
XX
PI Maisompierre PC, Masiakowski P, Yancopoulos GD;
XX
DR WPI; 1999-044584/04.
DR N-PSDB; AAV70207.
XX
PT DNA encoding receptor tyrosine kinase proteins - and corresponding
PT proteins
XX
PS Example; Fig 22; 194pp; English.
XX
CC The present invention describes nucleic acid molecules for ror-1,
CC ror-2, etk-1 and etk-2. Also described are the corresponding proteins:
CC Ror-1; Ror-2; Etk-1; and Etk-2. The proteins are orphan receptor
CC tyrosine kinases. The present sequence represents rat Etk-1.
XX
SQ Sequence 1005 AA;

Query Match 22.3%; Score 67; DB 20; Length 1005;
Best Local Similarity 31.6%; Pred. No. 22;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDCM-----DCASCR-----ARPHSDFCLGCAAAP 45
Db 273 ppkmhcsaegewlvpgkcmckagyeekngtcqvcprpffkasphsqtcskc---pp 326

RESULT 12
AAW29189
ID AAW29189 standard; Protein; 187 AA.

XX
AC AAY29189;
XX
DT 25-OCT-1999 (first entry)
XX
DE Amino acid sequence of a virulence factor encoded by ORF24368.
XX
KW Human pathogen; virulence polypeptide; virulence factor;
KW pathogenic infection; Pseudomonas aeruginosa infection.
XX
OS Pseudomonas aeruginosa.
XX
PN W09927129-A1.
XX
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25247.
XX
PR 25-NOV-1997; 97US-0066517.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;
PI Rahme LG, Tan M, Tsongalis J;
XX
DR WPI; 1999-357851/30.
XX
PT Virulence factors useful in developing disease treatments
XX
PS Disclosure; Fig 3; 228pp; English.
XX
CC The present sequence represents a Pseudomonas aeruginosa polypeptide
CC sequence. P. aeruginosa is an opportunistic human pathogen present in
CC soil water and plants. The specification describes virulence polypeptides
CC and nucleic acid sequence encoding such polypeptides. These sequences
CC can be used to identify a compound which is capable of decreasing the
CC expression of a pathogenic virulence factor. Compounds that inhibit
CC the expression or activity of virulence factor polypeptides can be
CC used to treat pathogenic infections, especially where the infection
CC is a P. aeruginosa infection.
CC note: the sequences given in the specification were poorly legible, and
CC in some instances assumptions were made as to the identity of the
CC residue; it is therefore possible that the sequence given below is
CC not entirely correct.
XX
SQ Sequence 187 AA;

Query Match 21.8%; Score 65.5; DB 20; Length 187;
Best Local Similarity 39.5%; Pred. No. 6.2;
Matches 17; Conservative 2; Mismatches 11; Indels 13; Gaps 3;

QY 6 TAPCSRGSSWSADLDCMDCASC-RARPHSDFCLGCAAAP 47
Db 123 sagecvpgpgwsg-----tgcapcsgrp-----caapppsp 153

RESULT 13
AAW12612
ID AAW12612 standard; Protein; 77 AA.
XX
AC AAW12612;
XX
DT 22-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO: 277 from WO 9906553.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
```





|    |   |   |
|----|---|---|
| PH | Key   | Location/Qualifiers                     |
| FT | Modified-site   | 83                                      |
| FT |   | /note= "potential phosphorylation site" |
| XX | W02000050588-A2.  |   |
| XX | 31-AUG-2000.  |   |
| XX | 01-FEB-2000; 2000WO-US02595.  |   |
| XX | 22-FEB-1999; 99US-0255381.  |   |
| XX | (INCY-) INCYTE PHARM INC.   |   |
| XX | Walker MG, Volkmut W, Klingler TM, Lal P;                                 |   |
| XX | WPI; 2000-558397/51.  |   |
| XX | Novel gene associated with colon cancer and coexpressed with one or       |   |
| PT | more known colon cancer genes in a number of biological samples, for      |   |
| PT | use in gene therapy   |   |
| XX | Claim 4; Page 32-33; 33pp; English.                                       |   |
| XX | The present sequence is encoded by a novel gene which is associated with  |   |
| CC | colon disease. The gene is coexpressed with one or more colon cancer      |   |
| CC | genes selected from carbonic anhydrase I, II and IV (CA I, II and IV),    |   |
| CC | carcinoembryonic antigen family of proteins (cea), colorectal carcinoma   |   |
| CC | tumour associated antigen (CO-029), down-regulated in adenoma (dra),      |   |
| CC | fatty-acid binding protein (fabp), galectin (galec), glutathione          |   |
| CC | peroxidase (gp2), guanylin (guan), cytokeratin 8 and 20 (ker 8 and 20),   |   |
| CC | cadherin (cadher) or intestinal mucin (muc-2). The present polynucleotide |   |
| CC | sequence is useful for diagnosing a disease or condition associated       |   |
| CC | with altered expression of the gene. The polynucleotide and encoded       |   |
| CC | polypeptides are useful for treating or preventing such diseases. They    |   |
| CC | are useful for diagnosing colon cancer, metastatic colon cancer, atrophic |   |
| CC | gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and   |   |
| CC | ulcerative colitis.   |   |
| XX | Sequence  | 115 AA;                                 |
| SQ |   | SQ                                      |

```

Query Match      21.3%  Score 64;  DB 21;  Length 115;
Best Local Similarity 30.9%;  Pred. No. 5.;
Matches 16;  Conservative 3;  Mismatches 23;  Indels 10;  Gaps 2;

QY  2  QAP-----GAPCSCGSSWSADLDKCM-DCASCRRARPSDFCLGCAAA 43
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  4  qapvvvtqpgvgpgpqpqpqpwnatgmcdcfdsdcqvcclcgftfcfclqagcva 55

```

Search completed: March 11, 2002, 15:14:14  
Job time: 98 sec

THIS PAGE BLANK (USPTO)



Db 273 PPKMHCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

## RESULT 2

US-08-440-815-2

; Sequence 2, Application US/08440815

; Patent No. 5798448

; GENERAL INFORMATION:

; APPLICANT: Caras, Ingrid W.

; APPLICANT: Winslow, John W.

; TITLE OF INVENTION: AL-1 Neurotrophic Factor

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/440,815

; FILING DATE: 15-MAY-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/330128

; FILING DATE: 27-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Torchia, Timothy E.

; REGISTRATION NUMBER: 36,700

; REFERENCE/DOCKET NUMBER: 920C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/952-8674

; TELEFAX: 415/952-8674

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 928 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-440-815-2

Query Match 22.3%; Score 67; DB 1; Length 928;

Best Local Similarity 31.6%; Pred. No. 6.6;

Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45

Db 273 PPKMHCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

## RESULT 3

US-08-486-449-2

; Sequence 2, Application US/08486449

; Patent No. 6280732

; GENERAL INFORMATION:

; APPLICANT: Caras, Ingrid W.

; APPLICANT: Winslow, John W.

; TITLE OF INVENTION: AL-1 Neurotrophic Factor

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,449

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/330128

; FILING DATE: 27-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Torchia, Timothy E.

; REGISTRATION NUMBER: 36,700

; REFERENCE/DOCKET NUMBER: P0920P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-8674

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 928 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-486-449-2

Query Match 22.3%; Score 67; DB 4; Length 928;

Best Local Similarity 31.6%; Pred. No. 6.6;

Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45

Db 273 PPKMHCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

## RESULT 4

US-08-469-537A-103

; Sequence 103, Application US/08469537A

; Patent No. 5843749

; GENERAL INFORMATION:

; APPLICANT: Maisonnier, et al.

; TITLE OF INVENTION: EHK AND ROR TYROSINE

; TITLE OF INVENTION: KINASES

; NUMBER OF SEQUENCES: 107

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

; STREET: 777 Old Saw Mill River Road

; CITY: Tarrytown

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10591

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,537A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USN 08/406,247

; FILING DATE: 17-MAR-1995

; APPLICATION NUMBER: USN 08/144,992

; FILING DATE: 28-OCT-1993

; APPLICATION NUMBER: USN 07/736,559

; FILING DATE: 26-JUL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Kempner, Ph.D., Gail M

; REGISTRATION NUMBER: 32,143

; REFERENCE/DOCKET NUMBER: REG 070C

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1005 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-537A-103

Query Match 22.3%; Score 67; DB 2; Length 1005;  
Best Local Similarity 31.6%; Pred. No. 7.1;  
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLKCM-----DCASCR-----ARPHSDFCGCAAAP 45  
DB 273 PPKMCSABGEWLVPIGKCMKAGYEKNGTCQVCPGPFKASPHSQTCSC-PP 326

## RESULT 5

US-08-336-708A-9  
Sequence 9, Application US/08336708A  
Patent No. 5521295

GENERAL INFORMATION:  
APPLICANT: Pacifici, Robert E.  
APPLICANT: Thomason, Arlen R.  
APPLICANT: Chang, Ming-Shi  
TITLE OF INVENTION: Hybrid Receptor Molecules  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,708A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oleski, Nancy  
REFERENCE/DOCKET NUMBER: A-241A  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 644 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-336-708A-9

Query Match 20.8%; Score 62.5; DB 1; Length 644;  
Best Local Similarity 34.0%; Pred. No. 15;  
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;

QY 9 CSRGSSWSADLKCM-----CAS-----CRARPHSDFCG-CAAAP 47  
DB 194 CPNGSCWAGAGEENCKLTKIIICACGRCGRKSPSDCHNCAAGCTGP 243

## RESULT 6

US-08-484-438-7  
Sequence 7, Application US/08484438  
Patent No. 5811098

Patent No. 5811098 5780031  
GENERAL INFORMATION:  
APPLICANT: Plowman, Gregory D.  
APPLICANT: Cuiuscou, Jean-Michel  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Siegall, Clay B.  
APPLICANT: Hellstr m, Ingegerd  
APPLICANT: Hellstr m, Karl E.  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1210 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-438-7

Query Match 20.8%; Score 62.5; DB 2; Length 1210;  
Best Local Similarity 34.0%; Pred. No. 28;  
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;

QY 9 CSRGSSWSADLKCM-----CAS-----CRARPHSDFCG-CAAAP 47  
DB 194 CPNGSCWAGAGEENCKLTKIIICACGRCGRKSPSDCHNCAAGCTGP 243

## RESULT 7

US-08-475-035-4  
Sequence 4, Application US/08475035  
Patent No. 5985553  
GENERAL INFORMATION:  
APPLICANT: KING, C. R.  
APPLICANT: KRAUS, MATTHIAS H.  
APPLICANT: AARONSON, STUART A.  
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM  
TITLE OF INVENTION: EGF RECEPTOR GENE

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,035  
FILING DATE: 7 Jun 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perryman, David G.  
REGISTRATION NUMBER: 33,438  
REFERENCE/DOCKET NUMBER: 1414.656  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-475-035-4

Query Match 20.8%; Score 62.5; DB 2; Length 1210;  
Best Local Similarity 34.0%; Pred. No. 28;  
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;  
QY 9 CSRGSSWSADLKDMD-----CAS-----CRARPHSDFLG-CAAAAPPAP 47  
DB 194 CPNGSCWAGGENCECKLTKIICACQCSGRCRGKSPSCCHNQCAAGCTGP 243

RESULT 8  
PCT-US95-02251-3  
Sequence 3, Application PC/TUS9502251  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02251  
FILING DATE: CONCURRENTLY HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/316,650  
FILING DATE: 30-SEP-1994  
CLASSIFICATION:  
APPLICATION NUMBER: US 08/199,780  
FILING DATE: 18-FEB-1994

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UMIC009P--  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02251-3

Query Match 20.8%; Score 62.5; DB 5; Length 1251;  
Best Local Similarity 29.4%; Pred. No. 29;  
Matches 15; Conservative 9; Mismatches 12; Indels 15; Gaps 3;

QY 9 CSRGSSWS-----ADLDKCMDCASCRRPHSDFC-----LGCAAAPPAPF 48  
DB 1000 CKQGFYDGNLLECDVDDECLDESNCR-----NGVCENTWRLPCACTPPAY 1046

RESULT 9  
US-08-199-780-3  
Sequence 3, Application US/08199780  
Patent No. 5763416  
GENERAL INFORMATION:  
APPLICANT: Bonadio, Jeffrey  
TITLE OF INVENTION: Gene Transfer Into Bone Cells  
TITLE OF INVENTION: And Tissues  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/199,780  
FILING DATE: 18-FEB-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UMIC:002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 320-7200  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1252 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-199-780-3

Query Match 20.8%; Score 62.5; DB 1; Length 1252;  
Best Local Similarity 29.4%; Pred. No. 29;  
Matches 15; Conservative 9; Mismatches 12; Indels 15; Gaps 3;

RESULT 11  
US-08-460-309-5  
; Sequence 5, Application US/08460309  
; Patent No. 5837496  
; GENERAL INFORMATION:  
; APPLICANT: Engvall, Eva  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; TITLE OF INVENTION: Fragments and Uses Thereof  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:

```

; Patent No. 5872231
; Patent No. 5872231 5840863
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; TITLE OF INVENTION: Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:

```

;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/125,077  
;/ FILING DATE: 22-SEP-1993  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US PCT/US 94/10730  
;/ FILING DATE: 21-SEP-1994  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US 07/472,319  
;/ FILING DATE: 30-JAN-1990  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US 07/919,951  
;/ FILING DATE: 27-JUL-1992  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Campbell, Cathryn A.  
;/ REGISTRATION NUMBER: 31,815  
;/ REFERENCE/DOCKET NUMBER: P-LA 9721  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (619) 535-9001  
;/ TELEFAX: (619) 535-8949  
;/ INFORMATION FOR SEQ ID NO: 5:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 3075 amino acids  
;/ TYPE: amino acid  
;/ TOPOLOGY: linear  
;/ US-08-125-077-5

Query Match 20.8%; Score 62.5; DB 2; Length 3075;  
Best Local Similarity 23.5%; Pred. No. 73;  
Matches 16; Conservative 6; Mismatches 23; Indels 23; Gaps 2;

QY 4 PGTAAPC-----SRGSSWSADLDKCMDCASCRCRAPHSDFC-----LGC 40  
| | | | | : : : : : | | | | |  
Db 1395 PLVAPCVPCSNHSDTCDPNTGKLCNGDNTAGDHCDVCTGYGKVTGSASDCALCAC 1454

QY 41 AAAPPAPF 48  
: | | | |

Db 1455 PHSPPASF 1462

RESULT 13  
US-08-976-255-11  
;/ Sequence 11, Application US/08976255  
;/ Patent No. 6136581  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Jono, Keith E.  
;/ APPLICANT: Plowman, Gregory  
;/ TITLE OF INVENTION: KINASE GENES AND USES  
;/ NUMBER OF SEQUENCES: 53  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Lyon & Lyon  
;/ STREET: 633 West Fifth Street  
;/ STREET: Suite 4700  
;/ CITY: Los Angeles  
;/ STATE: California  
;/ COUNTRY: U.S.A.  
;/ ZIP: 90071-2066  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
;/ MEDIUM TYPE: storage  
;/ COMPUTER: IBM Compatible  
;/ OPERATING SYSTEM: IBM P.C. DOS 5.0  
;/ SOFTWARE: FastSeq for Windows 2.0  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/976,255  
;/ FILING DATE: NO. 6136581ember 21, 1997  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:

;/ APPLICATION NUMBER: 60/031,675  
;/ FILING DATE: NO. 6136581ember 22, 1996  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Warburg, Richard J.  
;/ REGISTRATION NUMBER: 32,327  
;/ REFERENCE/DOCKET NUMBER: 229/182  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (213) 489-1600  
;/ TELEFAX: (213) 955-0440  
;/ TELEX: 67-3510  
;/ INFORMATION FOR SEQ ID NO: 11:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 1384 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: Protein  
;/ US-08-976-255-11

Query Match 20.1%; Score 60.5; DB 4; Length 1384;  
Best Local Similarity 32.7%; Pred. No. 54;  
Matches 16; Conservative 4; Mismatches 18; Indels 11; Gaps 2;

QY 4 PGTAAPC-----SRGSSWSADLDKCMDCASCRCRAPHSDFCGLGCAAPPA 46  
| | | | | : | | | | : | | | | |  
Db 523 PGVVPVLSAHSPSLGSEYFIRLEEAAPAG-----HDPDCAGCAPSPPA 566

RESULT 14  
US-09-286-529-20  
;/ Sequence 20, Application US/09286529  
;/ Patent No. 6297367  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Catherine Tribouley  
;/ TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES  
;/ FILE REFERENCE: 1408.003/200130.439C1  
;/ CURRENT APPLICATION NUMBER: US/09/286,529  
;/ CURRENT FILING DATE: 1999-04-05  
;/ NUMBER OF SEQ ID NOS: 25  
;/ SOFTWARE: FastSeq for Windows Version 3.0  
;/ SEQ ID NO 20  
;/ LENGTH: 211  
;/ TYPE: PRT  
;/ ORGANISM: Homo sapien  
;/ US-09-286-529-20

Query Match 19.8%; Score 59.5; DB 4; Length 211;  
Best Local Similarity 25.6%; Pred. No. 10;  
Matches 21; Conservative 6; Mismatches 14; Indels 41; Gaps 5;

QY 4 PGT---APCSRGS-----WSADLDKCMDC-----AS 27  
| | | | | : | | | | : | | | | |  
Db 54 PGTFVQRPCCRRSDPTTCGCPCCPRHYTFQWNY-LERCRCYNVLCGEREEARACHATHNRA 112  
| | | | | : | | | | : | | | | |  
QY 28 CRAR-----PHSDFCGLGCAAPP 45  
| | | | | : | | | | : | | | | |  
Db 113 CRCRTGFFAHAGFCLEHASCPP 134

RESULT 15  
US-09-286-529-17  
;/ Sequence 17, Application US/09286529  
;/ Patent No. 6297367  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Catherine Tribouley  
;/ TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES  
;/ FILE REFERENCE: 1408.003/200130.439C1  
;/ CURRENT APPLICATION NUMBER: US/09/286,529  
;/ CURRENT FILING DATE: 1999-04-05  
;/ NUMBER OF SEQ ID NOS: 25  
;/ SOFTWARE: FastSeq for Windows Version 3.0



```
; SEQ ID NO 17
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-286-529-17
```

```
Query Match      19.8%; Score 59.5; DB 4; Length 299;
Best Local Similarity 25.6%; Pred. No. 14;
Matches 21; Conservative 6; Mismatches 14; Indels 41; Gaps 5;

QY  4 PGT---APCSRGS-----WSADLDKCMDC-----AS 27
    ||| ||| |
Db  54 PGTFFQPCRRDPTTCGPPRHYTFWNY-LERCRCNVLCGEREEARACHATHNRA 112
    ||| ||| |
QY  28 CRAR---PHSDFCLGCAAPP 45
    ||| |
Db  113 CRCTGFFAHAGFCLEHASCPP 134
    ||| |
```

```
Search completed: March 11, 2002, 15:13:14
Job time: 38 sec
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:51 ; Search time 31.92 Seconds  
 (without alignments)  
 124.094 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_79  
 Perfect score: 301  
 Sequence: 1 EQAPGTAPCSRGSSWSADLD.....HSDFCIGCAAPAPFRLIW 52

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 73.5  | 24.4        | 261    | 2 G69099 | probable pyruvate  |
| 2          | 67    | 22.3        | 893    | 2 S51603 | receptor-like tyro |
| 3          | 67    | 22.3        | 898    | 2 S47489 | receptor tyrosine  |
| 4          | 67    | 22.3        | 981    | 2 S51604 | receptor-like tyro |
| 5          | 67    | 22.3        | 1005   | 2 S49015 | receptor tyrosine  |
| 6          | 66    | 21.9        | 549    | 2 B83085 | conserved hypotet  |
| 7          | 65.5  | 21.8        | 1013   | 2 I50615 | receptor-type prot |
| 8          | 64    | 21.3        | 1210   | 2 A53183 | epidermal growth f |
| 9          | 64    | 21.3        | 2871   | 2 A55624 | fibillin-1 precu   |
| 10         | 62.5  | 20.8        | 1210   | 1 GQHUE  | epidermal growth f |
| 11         | 62.5  | 20.8        | 1251   | 2 A57293 | latent transformin |
| 12         | 62.5  | 20.8        | 3075   | 2 S14458 | laminin alpha-1 ch |
| 13         | 62    | 20.6        | 2318   | 2 S45306 | notch 3 protein -  |
| 14         | 61    | 20.3        | 55     | 1 FEPE   | ferredoxin 2[4Fe-4 |
| 15         | 61    | 20.3        | 66     | 2 S58086 | metallothionein 3  |
| 16         | 61    | 20.3        | 68     | 2 A46034 | metallothionein 3, |
| 17         | 61    | 20.3        | 68     | 2 I67866 | growth inhibitory  |
| 18         | 61    | 20.3        | 112    | 2 T29620 | hypothetical prote |
| 19         | 61    | 20.3        | 390    | 2 F69086 | pyruvate formate-1 |
| 20         | 60.5  | 20.1        | 76     | 2 T03860 | TA20 protein - com |
| 21         | 60.5  | 20.1        | 414    | 2 T15947 | hypothetical prote |
| 22         | 60.5  | 20.1        | 644    | 2 A36325 | epidermal growth f |
| 23         | 60.5  | 20.1        | 1207   | 2 T00378 | KIAA0641 protein - |
| 24         | 60.5  | 20.1        | 1615   | 2 JE0372 | low density lipopr |
| 25         | 60    | 19.9        | 799    | 2 F83456 | xanthine dehydrog  |
| 26         | 60    | 19.9        | 837    | 1 A29312 | LDL receptor precu |
| 27         | 60    | 19.9        | 1548   | 2 S34583 | serine proteinase  |
| 28         | 59.5  | 19.8        | 431    | 1 B69092 | conserved hypotet  |
| 29         | 59    | 19.6        | 68     | 2 B46034 | metallothionein 3, |

#### ALIGNMENTS

##### RESULT 1

G69099

probable pyruvate synthase (EC 1.2.7.1) gamma chain - Methanobacterium thermoautotrop

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-May-2000

C:Accession: G69099

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: G69099

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <MTH>

A:Cross-references: GB:A6000929; GB:A6000666; NID:g2622853; PIDN:AA886210.1; PID:g262

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1740

C:Superfamily: pyruvate synthase gamma chain

C:Keywords: coenzyme A; oxidoreductase

Query Match 24.4%; Score 73.5; DB 2; Length 261;  
 Best Local Similarity 29.2%; Pred. No. 0.69;  
 Matches 19; Conservative 8; Mismatches 21; Indels 17; Gaps 4;

Qy 4 PGTAPCSRGSSWSA-----DLDKCMDCASCA-----RPHS---DFCLGCA-AAPPA 46

Db 191 PGSTVKNKTSWRTPKPVLDKDKCIDCDNCILFCPEGCINREHEIDYDYCKGCGCAEKC 250

Qy 47 PFRLL 51

Db 251 PVKAI 255

##### RESULT 2

S51603

receptor-like tyrosine kinase Etk-1 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999

C:Accession: S51603

R:Maisonnier, P.C.; Barrezuela, N.X.; Yancopoulos, G.D.

Oncogene 8, 3277-3288, 1993

A:Title: Etk-1 and Etk-2: two novel members of the Eph receptor-like tyrosine kinase

A:Reference number: S49015; MUID:94067777

A:Accession: S51603

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-893 <MAI>

A:Cross-references: EMBL:S68028

A>Note: the authors translated the codon GAC for residue 170 as Glu  
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat  
C:Keywords: ATP; transmembrane protein  
F:563-829/Domain: protein kinase homology <KIN>  
F:571-579/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 893;  
Best Local Similarity 31.6%; Pred. No. 8.4;  
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAFCRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45  
Db 273 PPKMHCSAEGEVLVPIGCKMCKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 3  
S47489  
receptor tyrosine kinase - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
C:Accession: S47489  
R:Taylor, V.; Pfarr, S.; Miescher, G.C.; Honegger, P.; Breitschopf, H.; Lassmann, H.; St  
submitted to the EMBL Data Library, April 1994  
A:Description: Expression and developmental regulation of EHK-1, a neuronal ELK-like rec  
A:Reference number: S47489  
A:Accession: S47489  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-898 <TAY>  
A:Cross-references: EMBL:X78689; NID:g531543; PIDN:CAA55357.1; PID:g531544  
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat  
C:Keywords: ATP; transmembrane protein  
F:568-834/Domain: protein kinase homology <KIN>  
F:576-584/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 898;  
Best Local Similarity 31.6%; Pred. No. 8.4;  
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAFCRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45  
Db 278 PPKMHCSAEGEVLVPIGCKMCKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 331

RESULT 4  
S51604  
receptor-like tyrosine kinase Ehk-1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 29-May-1998  
C:Accession: S51604  
R:Naisonnier, P.C.; Barrezaeta, N.X.; Yancopoulos, G.D.  
Oncogene 8, 3277-3288, 1993  
A:Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase fam  
A:Reference number: S49015; MUID:94067777  
A:Accession: S51604  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-981 <MAI>  
A:Cross-references: EMBL:S68029  
A>Note: the authors translated the codon GAC for residue 170 as Glu  
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat  
C:Keywords: ATP; transmembrane protein  
F:651-917/Domain: protein kinase homology <KIN>  
F:659-667/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 981;  
Best Local Similarity 31.6%; Pred. No. 9;  
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAFCRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45

Db 273 PPKMHCSAEGEVLVPIGCKMCKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 5  
S49015  
receptor tyrosine kinase Ehk-1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999  
C:Accession: S49015; S51602  
R:Naisonnier, P.C.; Barrezaeta, N.X.; Yancopoulos, G.D.  
Oncogene 8, 3277-3288, 1993  
A:Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase  
A:Reference number: S49015; MUID:94067777  
A:Accession: S49015  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1005 <MAI>  
A:Cross-references: EMBL:S68024  
A>Note: the authors translated the codon GAC for residue 170 as Glu  
A:Accession: S51602  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-305, 'G', 359-1005 <MA2>  
A:Cross-references: EMBL:S68026  
A>Note: the authors translated the codon GAC for residue 170 as Glu  
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea  
C:Keywords: ATP; transmembrane protein  
F:675-941/Domain: protein kinase homology <KIN>  
F:683-691/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 1005;  
Best Local Similarity 31.6%; Pred. No. 9.1;  
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAFCRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45  
Db 273 PPKMHCSAEGEVLVPIGCKMCKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 6  
E83085  
conserved hypothetical protein PA4488 [imported] - Pseudomonas aeruginosa (strain PAO  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: E83085  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.N.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: A82950; MUID:20437337  
A:Accession: E83085  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-549 <STO>  
A:Cross-references: GB:AE004862; GB:AE004091; NID:g9950716; PIDN:AAG07876.1; GSPDB:GN  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA4488  
C:Superfamily: Escherichia coli hypothetical protein b2226

Query Match 21.9%; Score 66; DB 2; Length 549;  
Best Local Similarity 36.7%; Pred. No. 7.5;  
Matches 18; Conservative 3; Mismatches 16; Indels 12; Gaps 2;

QY 4 PGTAFCRGSSWSADLDKCMDCASCRPHSDFCLGCAAAAPPFRLLW 52  
Db 380 PASAPTRRTAAWSADL----VLGATVNVHSD-----QPGPARLSW 416

## RESULT 7

A;Molecule type: mRNA  
A;Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A;Cross-references: EMBL:Z12608  
R;Heisermann, G.J.; Gill, G.N.  
J. Biol. Chem. 263, 13152-13158, 1988  
A;Title: Epidermal growth factor receptor threonine and serine residues phosphorylate  
R;Stever, D.A.; Verderame, M.F.  
Gene 148, 219-226, 1994  
A;Title: Identification of a complete Cdk7 receptor protein tyrosine kinase coding sequence  
A;Reference number: I50614; MUID:95047429  
A;Accession: I50615  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1013 <SIE>  
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60612.1; PID:g555618  
A;Accession: I50616  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-971, 'K', 973-1210 <VER>  
A;Cross-references: EMBL:X78987; NID:g488830; PIDN:CAA55587.1; PID:g488831  
R;Paria, B.C.; Das, S.K.; Andrews, G.K.; Dev, S.K.  
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
A;Title: Expression of the epidermal growth factor receptor gene is regulated in mouse  
A;Reference number: I49643; MUID:93126380  
A;Accession: I49643  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 12-20, 22-132 <RES>  
A;Cross-references: GB:L06864; NID:g193001; PIDN:AAA53029.1; PID:g567201  
C;Genetics:  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: Atg; growth factor receptor; kinase-related transforming protein; phospho  
F;649-915/Domain: protein kinase homology <KIN>  
F;657-665/Region: protein kinase ATP-binding motif  
F;938-1004/Domain: SAM homology <SAM>

Query Match 21.88; Score 65.5; DB 2; Length 1013;

Best Local Similarity 29.2%; Pred. No. 13;

Matches 19; Conservative 5; Mismatches 22; Indels 19; Gaps 4;

## QY 1

1 EQAPTAPCSRGSSWSADLCKM-----DCASCR-----ARPHSDPCLGCAAP 45

Db 245 DEAP-KHCSAEGEWLPIGKCLCKAGYEKNNTQVCVRPGFFKASPHSPSCSKC---PP 300

QY 46 APFRL 50

: |

Db 301 HSYTL 305

## RESULT 8

A;Molecule type: mRNA  
A;Residues: 1-1210 <LUE>  
A;Cross-references: GB:U03425  
R;Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.  
Oncogene 6, 673-676, 1991  
A;Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site  
A;Reference number: A43818; MUID:91232866  
A;Accession: A43818  
A;Molecule type: mRNA  
A;Residues: 1-714 <AVI>  
A;Cross-references: GB:X59698  
R;Eisinger, D.P.; Serrero, G.  
submitted to the EMBL Data Library, June 1992  
A;Reference number: S24942  
A;Accession: S24942

epidermal growth factor receptor precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999

C;Accession: A53183; A43818; S24942; A28941; S45325; I49643

R;Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;

Genes Dev. 8, 399-413, 1994

A;Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor

A;Reference number: A53183; MUID:94170986

A;Accession: A53183

A;Molecule type: mRNA

A;Residues: 1-1210 <LUE>

A;Cross-references: GB:U03425

R;Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.

Oncogene 6, 673-676, 1991

A;Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site

A;Reference number: A43818; MUID:91232866

A;Accession: A43818

A;Molecule type: mRNA

A;Residues: 1-714 <AVI>

A;Cross-references: GB:X59698

R;Eisinger, D.P.; Serrero, G.

submitted to the EMBL Data Library, June 1992

A;Reference number: S24942

A;Accession: S24942

Query Match 21.3%; Score 64; DB 2; Length 1210;

Best Local Similarity 31.7%; Pred. No. 22;

Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

## QY 4

4 PGTAP-----CSRGSSWSADLCKMND-----CAS-----CRARPHSDPCLG-CAAAP 47

Db 184 PSSCPKCDPSCPNCGSGGGEQCQLTKIICAQCSHRCGRSPSCDCHNQCAAGCTGP 243

## RESULT 9

A55624

fibrillin-1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 11-Jan-2000

C;Accession: A55624

R;Yin, W.; Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Pereira, L.; Ramire

J. Biol. Chem. 270, 1798-1806, 1995

A;Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin

A;Reference number: A55624; MUID:95130561

A;Accession: A55624

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2871 <YIN>

A;Cross-references: GB:L29454; NID:g575509; PIDN:AAA56840.1; PID:g575510

C;Genetics:

A;Gene: Fbn-1

C;Superfamily: unassigned EGF-related proteins; EGF homology

F;1201-1236/Domain: EGF homology <EGF>

[illegible]

R;Simmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
A;Reference number: A23062; MUID:85046483  
A;Accession: A23062  
A:Molecule type: mRNA  
A;Residues: 1028-1210 <SIM>  
R;Weber, W.; Gill, G.N.; Speiss, J.  
Science 224, 294-297, 1984  
A;Reference number: A05281; MUID:84172183  
A;Accession: A05281  
A:Molecule type: protein  
A;Residues: 25-30, S', 32-51;454-467 <WEB>  
R;Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.  
J. Biol. Chem. 260, 5205-5208, 1985  
A;Title: Identification of residues in the nucleotide binding site of the epidermal g  
A;Reference number: A60143; MUID:85182650  
A;Accession: A60143  
A:Molecule type: protein  
A;Residues: 740-744, X', 746-747 <RUS>  
R;Mroczkowski, B.; Mosig, G.; Cohen, S.  
Nature 309, 270-273, 1984  
A;Title: ATP-stimulated interaction between epidermal growth factor receptor and supe  
A;Reference number: A38023; MUID:84191554  
A;Contents: annotation; receptor activity  
A;Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
R;Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C  
Cell 59, 33-43, 1989  
A;Title: Functional independence of the epidermal growth factor receptor from a domai  
A;Reference number: A33331; MUID:90003233  
A;Contents: annotation; internalization signal  
C;Comment: Binding of EGF to the receptor leads to internalization of the EGF-recepto  
C;Genetics:  
A;Gene: GDB:EGFR  
A;Cross-references: GDB:120610; OMIM:131550  
A;Map position: 7p12.3-7p12.1  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-1210/Product: EGF receptor #status predicted <MAT>  
F;25-645/Domain: extracellular #status predicted <EXT>  
F;75-300/Domain: EGF receptor extracellular domain repeat <BE1>  
F;390-600/Domain: EGF receptor extracellular domain repeat <EE2>  
F;646-668/Domain: transmembrane #status predicted <TM>  
F;669-1210/Domain: intracellular #status predicted <INT>  
F;710-975/Domain: protein kinase homology <KIN>  
F;718-726/Region: protein kinase ATP-binding motif  
F;999-1046/Region: coated-pit mediated internalization signal  
F;1047-1210/Region: inhibitory  
F;128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status pre  
F;745/Active site: lys #status experimental

Query Match 20.8%; Score 62.5; DB 1; Length 1210;  
Best Local Similarity 34.0%; Pred. No. 32;  
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;

QY 9 CSRGSSASDLDKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47  
| | | | | | | | | | | | | | | | | | | |  
Db 194 CPNGSCHWAGENGCKLTKIICAOQCGRCKRKSPSCCHNQCAAGCTGP 243  
| | | | | | | | | | | | | | | | | | | |

RESULT 11  
A57293  
latent transforming growth factor beta-binding protein 3 precursor - mouse  
N;Alternate names: mitosis-inhibitory peptide  
C;Species: Mus musculus (house mouse)  
C;Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 11-Jan-2000  
C;Accession: A57293; A60487  
R;Yin, W.; Smiley, E.; Germiller, J.; Mechem, R.P.; Florer, J.B.; Wenstrup, R.J.; Bon  
J. Biol. Chem. 270, 10147-10160, 1995  
A;Title: Isolation of a novel latent transforming growth factor-beta binding protein  
A;Reference number: A57293; MUID:95247723  
A;Accession: A57293







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:50 ; Search time 19.78 Seconds  
(without alignments)  
96.389 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_79  
Perfect score: 301  
Sequence: 1 EQAPCTAPCSRGSSNSADLD.....HSDFCLGCAAAAPPAPFRLWL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1          | 73.5  | 24.4          | 81     | 1     | PORD_METTH  |
| 2          | 67    | 22.3          | 1005   | 1     | EPAS5_RAT   |
| 3          | 65.5  | 21.8          | 1013   | 1     | EPAS5_CHICK |
| 4          | 65    | 21.6          | 314    | 1     | Y127_HUMAN  |
| 5          | 64    | 21.3          | 1210   | 1     | EGFR_MOUSE  |
| 6          | 64    | 21.3          | 2871   | 1     | FBN1_MOUSE  |
| 7          | 62.5  | 20.8          | 1210   | 1     | EGFR_HUMAN  |
| 8          | 62.5  | 20.8          | 3075   | 1     | LMAL_HUMAN  |
| 9          | 62    | 20.6          | 54     | 1     | PER_PEPAS   |
| 10         | 62    | 20.6          | 2318   | 1     | NTC3_MOUSE  |
| 11         | 61.5  | 20.4          | 495    | 1     | MLP2_DROME  |
| 12         | 61    | 20.3          | 66     | 1     | MT3_RAT     |
| 13         | 61    | 20.3          | 68     | 1     | MT3_MOUSE   |
| 14         | 60    | 19.9          | 837    | 1     | LDLR_RABIT  |
| 15         | 60    | 19.9          | 1877   | 1     | PKG5_MOUSE  |
| 16         | 59.5  | 19.8          | 431    | 1     | YG84_METTH  |
| 17         | 59.5  | 19.8          | 547    | 1     | Z213_HUMAN  |
| 18         | 59    | 19.6          | 68     | 1     | MT3_HUMAN   |
| 19         | 59    | 19.6          | 558    | 1     | GPCL_RAT    |
| 20         | 59    | 19.6          | 1037   | 1     | EPAS5_HUMAN |
| 21         | 58.5  | 19.4          | 1255   | 1     | ERB2_HUMAN  |
| 22         | 58    | 19.3          | 68     | 1     | MT3_BOVIN   |
| 23         | 58    | 19.3          | 422    | 1     | Y140_HUMAN  |
| 24         | 58    | 19.3          | 443    | 1     | FBL4_CRIGR  |
| 25         | 58    | 19.3          | 443    | 1     | FBL4_MOUSE  |
| 26         | 57.5  | 19.1          | 624    | 1     | NIFA_AZOLI  |
| 27         | 57.5  | 19.1          | 1696   | 1     | PKL5_BRACL  |
| 28         | 57.5  | 19.1          | 2820   | 1     | NFL_RAT     |
| 29         | 57.5  | 19.1          | 2839   | 1     | NFL_HUMAN   |
| 30         | 57.5  | 19.1          | 2841   | 1     | NFL_MOUSE   |
| 31         | 57.5  | 19.1          | 3133   | 1     | HMCT_BOMMO  |
| 32         | 57    | 18.9          | 983    | 1     | EPAS3_MOUSE |
| 33         | 56.5  | 18.8          | 316    | 1     | CC12_CABEL  |

|    |      |      |      |   |            |
|----|------|------|------|---|------------|
| 34 | 56.5 | 18.8 | 316  | 1 | CC13_CABEL |
| 35 | 56   | 18.6 | 215  | 1 | HP25_TAMAS |
| 36 | 56   | 18.6 | 684  | 1 | FBL1_CHICK |
| 37 | 56   | 18.6 | 1217 | 1 | EGF_MOUSE  |
| 38 | 55.5 | 18.4 | 59   | 1 | FER2_DESDN |
| 39 | 55.5 | 18.4 | 279  | 1 | SLI3_MOUSE |
| 40 | 55.5 | 18.4 | 279  | 1 | SLI3_RAT   |
| 41 | 55.5 | 18.4 | 1278 | 1 | NPC1_HUMAN |
| 42 | 55   | 18.3 | 252  | 1 | DHSB_BACSU |
| 43 | 55   | 18.3 | 394  | 1 | FDHB_METIF |
| 44 | 54.5 | 18.1 | 265  | 1 | YIT6_YEAST |
| 45 | 54.5 | 18.1 | 464  | 1 | YMG8_YEAST |

## ALIGNMENTS

| RESULT PORD_METTH | 1  | STANDARD;                         | PRT; | 81 AA. |
|-------------------|--|-----------------------------------|------|--------|
| AC                | P56815;  | 2000 (Rel. 39, Created)           |      |        |
| DT                | 30-MAY-2000  | (Rel. 39, Last sequence update)   |      |        |
| DT                | 30-MAY-2000  | (Rel. 40, Last annotation update) |      |        |
| DE                | PYRUVATE SYNTHASE SUBUNIT PORD (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE DELTA CHAIN) (POR) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE DELTA SUBUNIT).   |                                   |      |        |
| DE                | PORD OR MTH1740.1.   |                                   |      |        |
| GN                | Methanobacterium thermoautotrophicum.  |                                   |      |        |
| OS                | Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;   |                                   |      |        |
| OC                | Methanothermobacter.   |                                   |      |        |
| OX                | NCBI_TaxID=145262;   |                                   |      |        |
| RN                | [1]  |                                   |      |        |
| RP                | SEQUENCE FROM N.A.   |                                   |      |        |
| RC                | STRAIN=DELTA H;  |                                   |      |        |
| RX                | MEDLINE=98037514; PubMed=9371463;  |                                   |      |        |
| RA                | Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R., Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S., McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M., Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.; "Complete genome sequence of Methanobacterium thermoautotrophicum delta: functional analysis and comparative genomics."; |                                   |      |        |
| RL                | J. Bacteriol. 179:7135-7155(1997).   |                                   |      |        |
| CC                | -!- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN = ACETYL-COA + CO(2) + REDUCED FERREDOXIN.  |                                   |      |        |
| CC                | -!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS.   |                                   |      |        |
| CC                | -!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE GAMMA CHAIN.   |                                   |      |        |
| CC                | -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.   |                                   |      |        |
| CC                | -!- CAUTION: THERE SEEMS TO BE A SEQUENCING ERROR THAT FUSES TOGETHER PORC AND PORD. WE HAVE CUT THE ORF INTO ITS TWO CONSTITUENTS.  |                                   |      |        |
| CC                | -----  |                                   |      |        |
| CC                | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> )      |                                   |      |        |
| CC                | -----  |                                   |      |        |
| DR                | EMBL; AE000929; AAB86210.1; ALT_INIT.  |                                   |      |        |
| DR                | HSSP; P00195; ICLF.  |                                   |      |        |
| DR                | InterPro; IPR001450; 4FE4S_ferrdxin.   |                                   |      |        |
| DR                | Pfam; PF00037; fer4; 2.  |                                   |      |        |
| DR                | PRINTS; PR00353; 4FE4SFRDXIN   |                                   |      |        |
| DR                | PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.   |                                   |      |        |
| KW                | Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S; Complete proteome.  |                                   |      |        |
| KW                | METAL 34   |                                   |      |        |
| FT                | METAL 37   |                                   |      |        |
| FT                | METAL 40   |                                   |      |        |
| FT                | IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  |                                   |      |        |
| FT                | IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  |                                   |      |        |
| FT                | IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  |                                   |      |        |

```
FT METAL 44 44 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 60 60 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 63 63 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 66 66 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 70 70 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 81 AA; 9121 MW; 219A9CCA8A41604 CRC64;

Query Match 24.4%; Score 73.5; DB 1; Length 81;
Best Local Similarity 29.2%; Pred No. 0.072; 21; Indels 17; Gaps 4;
Matches 19; Conservative 8; Mismatches 21; Indels 17; Gaps 4;

QY 4 PGTAFCRSGSSWSA-----DLKCDKDCASCRA-----RPHS-----DFCLCA-AAPPA 46
DB 11 PGSTVKNTGSRTRKPVLDKDKICDDNCILFCPEGGINREHEIDYDYCKGGCAEKC 70
QY 47 PFRLL 51
DB 71 PVKAI 75

RESULT 2
EPA5_RAT
AC EPA5_RAT STANDARD; PRT: 1005 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE KINASE RECEPTOR EHK-1) (EPH HOMOLOG KINASE-1).
GN EPA5 OR EHK1 OR EHK-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5).
RX STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=94067777; PubMed=7504232;
RA Maisongier P.C., Barrezeueta N.X., Yancopoulos G.D.;
RT "Ehk-1 and EHK-2: two novel members of the Eph receptor-like tyrosine
RT kinase family with distinctive structures and neuronal expression.";
RL Oncogene 8:3277-3288(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RX STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=95206467; PubMed=7898646;
RA Taylor V., Miescher G.C., Pfarr S., Honegger P., Breitschopf H.,
RA Lassmann H., Steck A.J.;
RT "Expression and developmental regulation of EHK-1, a neuronal
RT Elk-like receptor tyrosine kinase in brain.";
RL Neuroscience 63:163-178(1994).
CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 6 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5 AND
CC 6, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC SYSTEM. PREDOMINANTLY EXPRESSED IN NEURONS.
CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
CC RECEPTOR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; X78589; CAA55357.1; -.
DR HSSP; P00523; 2PTK.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001090; Ephrin_rcptor.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001777; FN_III.
DR InterPro; IPR001426; Rceptor_tyr_kin_v.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001245; Tyr_kin.
DR Pfam; PF01404; EPH_lbd; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD001495; Ephrin_rcptor; 1.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 1005 EPHRIN TYPE-A RECEPTOR 5.
FT DOMAIN 27 575 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 576 596 POTENTIAL.
FT DOMAIN 597 1005 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 222 356 CVS-RICH.
FT DOMAIN 357 466 FIBRONECTIN TYPE-III 1.
FT DOMAIN 467 563 FIBRONECTIN TYPE-III 2.
FT DOMAIN 577 938 PROTEIN KINASE.
FT DOMAIN 965 1005 SAM.
FT BIND 683 691 ATP (BY SIMILARITY).
FT BINDING 709 709 ATP (BY SIMILARITY).
FT ACT_SITE 802 802 BY SIMILARITY.
FT MOD_RES 652 652 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 658 658 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 835 835 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 984 984 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 10 20 GRRTQGRGGG -> DADGPRAQASWCHARR (IN ISOFORM 6).
FT VARSPLIC 306 357 VCRPGEFKASPHSQTCKPPHSYTHEASTSCVCEKDYFR
FT RESDEPTWACT -> G (IN ISOFORM 2 AND ISOFORM 5).
FT VARSPLIC 358 470 RPPSAPRNAISNNVNETSVFLEWIPPADTGGGKDVSYILCK
FT KCSHAGVCECGEGLVLPQOIGLKNYSVMADPLAHNTY
FT TCEIHAVGVSDLSPTGYVSVNVTNQA -> T (IN ISOFORM 3, ISOFORM 4, ISOFORM 5 AND ISOFORM 6).
FT VARSPLIC 597 621 SGSCCECGGRASSICAVAHPSLIW -> R (IN ISOFORM 4 AND ISOFORM 5).
FT CONFLICT 170 170 D -> E (IN REF. 2).
FT CONFLICT 566 566 G -> A (IN REF. 2).
FT CONFLICT 578 578 G -> A (IN REF. 2).
FT CONFLICT 669 669 G -> A (IN REF. 2).
FT CONFLICT 708 708 T -> I (IN REF. 2).
FT CONFLICT 979 979 T -> I (IN REF. 2).
SQ SEQUENCE 1005 AA; 111007 MW; 1AED42C99693C574 CRC64;
```

```
Query Match      22.3%; Score 67; DB 1; Length 1005;
Best Local Similarity 31.6%; Pred. No. 3.2;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45
Db 273 PPKMCSAEGEWLVPIGRKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---pp 326

RESULT 3
EPAS_CHICK STANDARD; PRT; 1013 AA.
AC P54755; 007495;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE KINASE RECEPTOR CEK7).
GN EPHA5 OR CEK7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body wall;
RX MEDLINE=95047429; PubMed=7958948;
RA Slevier D.A., Verderame M.F.;
RT "Identification of a complete Cek7 receptor protein tyrosine kinase
RT coding sequence and cDNAs of alternatively spliced transcripts.";
RL Gene 148:219-226(1994).
RN [2]
RP SEQUENCE OF 512-1013 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93288394; PubMed=8510926;
RA Sajjadi F.G., Pasquale E.B.;
RT "Five novel avian Eph-related tyrosine kinases are differentially
RT expressed.";
RL Oncogene 8:1807-1813(1993).
CC -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC EPHRIN-A1, -A2, -A3, -A4 AND -A5. ABLE TO COLLAPSE GROWTH CONES.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1, 2 AND 3 (SHOWN
CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: DETECTED IN THE 10-DAY EMBRYONIC BRAIN, WEAKER
CC EXPRESSION IN THE REST OF THE 10-DAY EMBRYO. UNDETECTED IN ADULT
CC TISSUES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
CC RECEPTOR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03910; AAB60613.1; -
CC EMBL; U03910; AAB60614.1; -
CC EMBL; U03910; AAB60612.1; -
CC EMBL; Z19058; CAA79508.1; -
CC HSP; P00523; 2PTK
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001090; Ephrin_rcptor.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001777; FN_III.
CC InterPro; IPR001426; Rptor_tyr_kin_v.
```

```
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001245; Tyr_kin.
DR Pfam; PF01404; EPH_lbd; 1.
DR Pfam; PF00041; fn3_2
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD001495; Ephrin_rcptor; 1.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00060; FN3_2.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 1013 EPHRIN TYPE-A RECEPTOR 5.
FT DOMAIN 32 549 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 550 570 POTENTIAL.
FT DOMAIN 571 1013 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 196 330 CYS-RICH.
FT DOMAIN 331 440 FIBRONECTIN TYPE-III 1.
FT DOMAIN 441 537 FIBRONECTIN TYPE-III 2.
FT DOMAIN 651 912 PROTEIN KINASE.
FT DOMAIN 939 1013 SAM.
FT SITE 1011 1013 PDZ-BINDING MOTIF (POTENTIAL).
FT NP_BIND 657 665 ATP (BY SIMILARITY).
FT BINDING 683 683 ATP (BY SIMILARITY).
FT ACT_SITE 776 776 BY SIMILARITY.
FT MOD_RES 626 626 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 632 632 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 809 809 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 958 958 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPPLIC 280 443 MISSING (IN ISOFORM 2).
FT VARSPPLIC 573 595 SCCDHCGCGWASSLRVAYPSLIW -> R (IN ISOFORM 1 AND ISOFORM 2).
FT CONFLICT 981 1013 LRRGLVTLVGHQKIMNSIQEMKVLVNGMVPL -> ESPC
FT SEQUENCE 1013 AA; 112245 MW; AC36FDGAEBF38382 CRC64;
SQ
```

```
Query Match      21.8%; Score 65.5; DB 1; Length 1013;
Best Local Similarity 29.2%; Pred. No. 4.7;
Matches 19; Conservative 5; Mismatches 22; Indels 19; Gaps 4;

QY 1 EQAPGTCPSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45
Db 245 DEAP-KMHCASBEGEWLVPIGRKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---pp 300

QY 46 APFRL 50
Db 301 HSYTL 305
```

```
RESULT 4
Y127_HUMAN STANDARD; PRT; 314 AA.
ID Y127_HUMAN
AC Q14140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
```

DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN KIAA0127.  
GN KIAA0127.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=96127530; PubMed=8590280;  
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. IV.  
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by  
RT analysis of cDNA clones from human cell line KG-1.";  
RL DNA Res. 2:167-174(1995).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
KW Hypothetical protein.  
DR EMBL; D50917; BAA09476.1; -  
SQ SEQUENCE 314 AA; 33896 MW; E43107FC565AAC31 CRC64;  
  
Query Match 21.6%; Score 65; DB 1; Length 314;  
Best Local Similarity 33.3%; Pred. No. 2;  
Matches 16; Conservative 6; Mismatches 24; Indels 2; Gaps 1;  
  
QY 3 APTACPSRGSSWSDLCMDACSCARPHSDFCLGCAAPAPPPRL 50  
DB 112 SPSSHPCDLGS--TTPLACLTPASLLEDDDDTFC\*SQAMQPTAPRKL 157  
  
RESULT 5  
EGFR\_MOUSE STANDARD; PRT; 1210 AA.  
AC Q01279;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112).  
GN EGFR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RX MEDLINE=93026370; PubMed=1408137;  
RA Avivi A., Skorecki K., Yayon A., Givol D.;  
RT "Promoter region of the murine fibroblast growth factor receptor 2.  
RT (bek/KGFR) gene.";  
RL Oncogene 7:1957-1962(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C, AND CD-1; TISSUE=Liver, and Decidua;  
RX MEDLINE=93126380; PubMed=7678348;  
RA Paria B.C., Das S.K., Andrews G.K., Dey S.K.;  
RT "Expression of the epidermal growth factor receptor gene is regulated  
RT in mouse blastocysts during delayed implantation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RA Hibbs M.L.;  
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6/C3; TISSUE=Liver;  
RX MEDLINE=94170986; PubMed=8125255;  
RA Luetke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,  
RA Jenkins N.A., Lee D.C.;  
RT "The mouse waved-2 phenotype results from a point mutation in the EGF  
RT receptor tyrosine kinase.";  
RL Genes Dev. 8:399-413(1994).  
RN [5]  
RP SEQUENCE OF 1-714 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91232866; PubMed=2030916;  
RA Avivi A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;  
RT "Comparison of EGF receptor sequences as a guide to study the ligand  
RT binding site.";  
RL Oncogene 6:673-676(1991).  
RN [6]  
RP SEQUENCE OF 969-1117 FROM N.A.  
RC STRAIN=C3H;  
RA Eisinger D.P., Serrero G.;  
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF, AND  
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
CC PROTEIN TYROSINE PHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- MISCELLANEOUS: BINDING OF EGF TO THE RECEPTOR LEADS TO  
CC INTERNALIZATION OF THE EGF-RECEPTOR COMPLEX, INDUCTION OF THE  
CC TYROSINE KINASE ACTIVITY, STIMULATION OF CELL DNA SYNTHESIS, AND  
CC CELL PROLIFERATION.  
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; X78987; CAA55587.1; -  
EMBL; U03425; AAA17899.1; -  
EMBL; X59698; CAA42219.1; -  
EMBL; L06864; AAA53029.1; -  
EMBL; Z12608; CAA78249.1; -  
HSSP; P11362; IFGI.  
DR MGD; MGI:95294; Egfr.  
DR InterPro: IPR000494; EGFR\_L.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_kin.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR SMART; SM00261; FU; 3.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.  
FT DOMAIN 25 647 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 648 670 POTENTIAL.  
FT DOMAIN 671 1210 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 75 300 APPROXIMATE.  
FT REPEAT 390 600 APPROXIMATE.  
FT DOMAIN 1028 1071 SER-RICH.  
FT DOMAIN 714 981 PROTEIN KINASE.  
FT NP\_BIND 720 728 ATP (BY SIMILARITY).



|    |          |      |      |                               |
|----|----------|------|------|-------------------------------|
| FT | DOMAIN   | 2402 | 2443 | EGF-LIKE 41, CALCIUM-BINDING. |
| FT | DOMAIN   | 2444 | 2484 | EGF-LIKE 42, CALCIUM-BINDING. |
| FT | DOMAIN   | 2485 | 2523 | EGF-LIKE 43, CALCIUM-BINDING. |
| FT | DOMAIN   | 2524 | 2566 | EGF-LIKE 44, CALCIUM-BINDING. |
| FT | DOMAIN   | 2567 | 2606 | EGF-LIKE 45, CALCIUM-BINDING. |
| FT | DOMAIN   | 2607 | 2647 | EGF-LIKE 46, CALCIUM-BINDING. |
| FT | DOMAIN   | 2648 | 2687 | EGF-LIKE 47, CALCIUM-BINDING. |
| FT | DISULFID | 85   | 94   | BY SIMILARITY.                |
| FT | DISULFID | 89   | 100  | BY SIMILARITY.                |
| FT | DISULFID | 102  | 111  | BY SIMILARITY.                |
| FT | DISULFID | 119  | 129  | BY SIMILARITY.                |
| FT | DISULFID | 123  | 134  | BY SIMILARITY.                |
| FT | DISULFID | 136  | 145  | BY SIMILARITY.                |
| FT | DISULFID | 150  | 160  | BY SIMILARITY.                |
| FT | DISULFID | 154  | 166  | BY SIMILARITY.                |
| FT | DISULFID | 168  | 177  | BY SIMILARITY.                |
| FT | DISULFID | 230  | 262  | BY SIMILARITY.                |
| FT | DISULFID | 257  | 271  | BY SIMILARITY.                |
| FT | DISULFID | 273  | 286  | BY SIMILARITY.                |
| FT | DISULFID | 292  | 304  | BY SIMILARITY.                |
| FT | DISULFID | 299  | 313  | BY SIMILARITY.                |
| FT | DISULFID | 315  | 328  | BY SIMILARITY.                |
| FT | DISULFID | 433  | 465  | BY SIMILARITY.                |
| FT | DISULFID | 460  | 474  | BY SIMILARITY.                |
| FT | DISULFID | 476  | 488  | BY SIMILARITY.                |
| FT | DISULFID | 494  | 504  | BY SIMILARITY.                |
| FT | DISULFID | 499  | 513  | BY SIMILARITY.                |
| FT | DISULFID | 515  | 528  | BY SIMILARITY.                |
| FT | DISULFID | 534  | 546  | BY SIMILARITY.                |
| FT | DISULFID | 541  | 555  | BY SIMILARITY.                |
| FT | DISULFID | 557  | 570  | BY SIMILARITY.                |
| FT | DISULFID | 576  | 587  | BY SIMILARITY.                |
| FT | DISULFID | 582  | 596  | BY SIMILARITY.                |
| FT | DISULFID | 598  | 611  | BY SIMILARITY.                |
| FT | DISULFID | 617  | 628  | BY SIMILARITY.                |
| FT | DISULFID | 623  | 637  | BY SIMILARITY.                |
| FT | DISULFID | 639  | 652  | BY SIMILARITY.                |
| FT | DISULFID | 727  | 739  | BY SIMILARITY.                |
| FT | DISULFID | 734  | 748  | BY SIMILARITY.                |
| FT | DISULFID | 750  | 763  | BY SIMILARITY.                |
| FT | DISULFID | 769  | 781  | BY SIMILARITY.                |
| FT | DISULFID | 776  | 790  | BY SIMILARITY.                |
| FT | DISULFID | 792  | 805  | BY SIMILARITY.                |
| FT | DISULFID | 811  | 821  | BY SIMILARITY.                |
| FT | DISULFID | 816  | 830  | BY SIMILARITY.                |
| FT | DISULFID | 832  | 845  | BY SIMILARITY.                |
| FT | DISULFID | 914  | 926  | BY SIMILARITY.                |
| FT | DISULFID | 921  | 935  | BY SIMILARITY.                |
| FT | DISULFID | 937  | 950  | BY SIMILARITY.                |
| FT | DISULFID | 1032 | 1044 | BY SIMILARITY.                |
| FT | DISULFID | 1039 | 1053 | BY SIMILARITY.                |
| FT | DISULFID | 1055 | 1068 | BY SIMILARITY.                |
| FT | DISULFID | 1074 | 1086 | BY SIMILARITY.                |
| FT | DISULFID | 1081 | 1095 | BY SIMILARITY.                |
| FT | DISULFID | 1097 | 1111 | BY SIMILARITY.                |
| FT | DISULFID | 1117 | 1129 | BY SIMILARITY.                |
| FT | DISULFID | 1124 | 1138 | BY SIMILARITY.                |
| FT | DISULFID | 1140 | 1153 | BY SIMILARITY.                |
| FT | DISULFID | 1159 | 1171 | BY SIMILARITY.                |
| FT | DISULFID | 1166 | 1180 | BY SIMILARITY.                |
| FT | DISULFID | 1182 | 1195 | BY SIMILARITY.                |
| FT | DISULFID | 1201 | 1212 | BY SIMILARITY.                |
| FT | DISULFID | 1208 | 1221 | BY SIMILARITY.                |
| FT | DISULFID | 1223 | 1236 | BY SIMILARITY.                |
| FT | DISULFID | 1242 | 1254 | BY SIMILARITY.                |
| FT | DISULFID | 1249 | 1263 | BY SIMILARITY.                |
| FT | DISULFID | 1265 | 1278 | BY SIMILARITY.                |
| FT | DISULFID | 1284 | 1296 | BY SIMILARITY.                |
| FT | DISULFID | 1291 | 1305 | BY SIMILARITY.                |
| FT | DISULFID | 1307 | 1320 | BY SIMILARITY.                |
| FT | DISULFID | 1326 | 1339 | BY SIMILARITY.                |
| FT | DISULFID | 1333 | 1348 | BY SIMILARITY.                |
| FT | DISULFID | 1350 | 1361 | BY SIMILARITY.                |

|    |          |      |      |                |
|----|----------|------|------|----------------|
| FT | DISULFID | 1367 | 1380 | BY SIMILARITY. |
| FT | DISULFID | 1374 | 1389 | BY SIMILARITY. |
| FT | DISULFID | 1391 | 1402 | BY SIMILARITY. |
| FT | DISULFID | 1408 | 1420 | BY SIMILARITY. |
| FT | DISULFID | 1415 | 1429 | BY SIMILARITY. |
| FT | DISULFID | 1431 | 1444 | BY SIMILARITY. |
| FT | DISULFID | 1450 | 1461 | BY SIMILARITY. |
| FT | DISULFID | 1456 | 1470 | BY SIMILARITY. |
| FT | DISULFID | 1472 | 1485 | BY SIMILARITY. |
| FT | DISULFID | 1491 | 1502 | BY SIMILARITY. |
| FT | DISULFID | 1497 | 1511 | BY SIMILARITY. |
| FT | DISULFID | 1513 | 1526 | BY SIMILARITY. |
| FT | DISULFID | 1610 | 1622 | BY SIMILARITY. |
| FT | DISULFID | 1617 | 1631 | BY SIMILARITY. |
| FT | DISULFID | 1633 | 1646 | BY SIMILARITY. |
| FT | DISULFID | 1652 | 1663 | BY SIMILARITY. |
| FT | DISULFID | 1658 | 1672 | BY SIMILARITY. |

Query Match 21.3%; Score 64; DB 1; Length 2871;  
Best Local Similarity 26.1%; Pred. No. 17;  
Matches 18; Conservative 5; Mismatches 18; Indels 28; Gaps 2;

QY 9 CSRGSWSADLDKCMD-----CASCRRPHS-----DFCLGC 40  
| | | | | : | : | : | | |  
DB 2040 CPFGFSWSSGRCDLRMSYCYAKFEGGKCSKSRNHSKQCCCALKGEGWGPCELC 2099

QY 41 AAAPPAPER 49  
| | | | |  
DB 2100 PTEPDEAFR 2108

RESULT 7  
EGFR\_HUMAN  
ID EGFR\_HUMAN STANDARD; PRT; 1210 AA.  
AC P00533; P06268; Q14225;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112).  
GN EGFR OR ERBB1  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84219729; PubMed=6328312;  
RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,  
Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
expression of the amplified gene in A431 epidermoid carcinoma cells.";  
RL Nature 309:418-425(1984).  
RN [2]  
RP SEQUENCE OF 713-924 FROM N.A.  
RX MEDLINE=84196372; PubMed=6326261;  
RA Lin C.R., Chen W.S., Krueger W., Stolarsky L.S., Weber W.,  
Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
RT "Expression cloning of human EGF receptor complementary DNA: gene  
amplification and three related messenger RNA products in A431  
cells.";  
RL Science 224:843-848(1984).  
RN [3]  
RP SEQUENCE OF 150-962 FROM N.A.  
RX MEDLINE=84245835; PubMed=6330563;  
RA Xu Y., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
Roe B.A., Merlino G.T., Pastan I.;  
RT "Human epidermal growth factor receptor cDNA is homologous to a  
variety of RNAs overproduced in A431 carcinoma cells.";  
RL Nature 309:806-810(1984).  
RN [4]  
RP SEQUENCE OF 1028-1210 FROM N.A.  
RX MEDLINE=85046483; PubMed=6093780;

RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,  
RA O'Malley B.W.;  
RT "Isolation of an evolutionarily conserved epidermal growth factor  
RT receptor cDNA from human A431 carcinoma cells.";  
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
[5]  
RX SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE=88217333; PubMed=3329716;  
RA Haley J., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
RA Waterfield M.;  
RT "The human EGF receptor gene: structure of the 110 kb locus and  
RT identification of sequences regulating its transcription.";  
RL Oncogene Res. 1:375-396(1987).  
[6]  
RX SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE=91107677; PubMed=1988448;  
RA Haley J.D., Waterfield M.D.;  
RT "Contributory effects of de novo transcription and premature  
RT transcript termination in the regulation of human epidermal growth  
RT factor receptor proto-oncogene RNA synthesis.";  
RL J. Biol. Chem. 266:1746-1753(1991).  
[7]  
RX SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE=85270438; PubMed=2991899;  
RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
RT "Characterization and sequence of the promoter region of the human  
RT epidermal growth factor receptor gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
[8]  
RX SEQUENCE OF 540.  
RA Kohda D.;  
RL Submitted (SEP-1997) to the SWISS-PROT data bank.  
[9]  
RX RECEPTOR ACTIVITY.  
RX MEDLINE=84191554; PubMed=6325948;  
RA Mroczkowski B., Mosig G., Cohen S.;  
RT "ATP-stimulated interaction between epidermal growth factor receptor  
RT and supercoiled DNA.";  
RL Nature 309:270-273(1984).  
[10]  
RX PHOSPHORYLATION.  
RX MEDLINE=89278137; PubMed=2543678;  
RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
RA Hawk R., Givol D., Ullrich A., Schlessinger J.;  
RT "All autophosphorylation sites of epidermal growth factor (EGF)  
RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
RT Identification of a novel site in EGF receptor.";  
RL J. Biol. Chem. 264:10667-10671(1989).  
[11]  
RX REVIEW.  
RX MEDLINE=87297456; PubMed=3039909;  
RA Carpenter G.;  
RT "Receptors for epidermal growth factor and other polypeptide  
RT mitogens.";  
RL Annu. Rev. Biochem. 56:881-914(1987).  
CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,  
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
CC VACCINIA VIRUS GROWTH FACTOR.  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
CC PROTEIN TYROSINE PHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- MISCELLANEOUS: BINDING OF EGF TO THE RECEPTOR LEADS TO  
CC INTERNALIZATION OF THE EGF-RECEPTOR COMPLEX, INDUCTION OF THE  
CC TYROSINE KINASE ACTIVITY, STIMULATION OF CELL DNA SYNTHESIS, AND  
CC CELL PROLIFERATION.  
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; X00588; CAA25240.1; -;  
DR EMBL; X06370; CAA29668.1; -;  
DR EMBL; X00663; CAA25282.1; -;  
DR EMBL; M38425; AAA63171.1; -;  
DR EMBL; M11234; AAA52370.1; -;  
DR PIR; A00641; GQHUE;  
DR PIR; A00642; GQHUE2;  
DR PIR; A23062; A23062;  
DR HSP; P11362; IFGI.  
DR SWISS-2DPAGE; P00533; HUMAN.  
DR MIM; 131550; -;  
DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002174; Furin-like.  
DR InterPro; IPR001245; Tyr\_kin.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR SMART; SM00261; FU; 3.  
DR SMART; SM00219; Tykac; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Transmembrane; Glycoprotein; Repeat; Receptor; Signal; Transferase;  
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT SIGNAL 1 24  
FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.  
FT DOMAIN 25 645 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 646 668 POTENTIAL.  
FT DOMAIN 669 1210 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 75 300 APPROXIMATE.  
FT REPEAT 390 600 APPROXIMATE.  
FT DOMAIN 1025 1071 SER-RICH.  
FT DOMAIN 712 979 PROTEIN KINASE.  
FT NP\_BIND 718 726 ATP (BY SIMILARITY).  
FT BINDING 745 745 ATP (BY SIMILARITY).  
FT ACT\_SITE 837 837 BY SIMILARITY.  
FT MOD\_RES 678 678 PHOSPHORYLATION (BY PKC).  
FT MOD\_RES 1092 1092 PHOSPHORYLATION (AUTO-).  
FT MOD\_RES 1110 1110 PHOSPHORYLATION (AUTO-).  
FT MOD\_RES 1172 1172 PHOSPHORYLATION (AUTO-).  
FT MOD\_RES 1197 1197 PHOSPHORYLATION (AUTO-).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 540 540 N -> K (IN REF. 1).  
SQ SEQUENCE 1210 AA; 134277 MW; D8A2A50B4EFB6ED2 CRC64;

Query Match 20.8%; Score 62.5; DB 1; Length 1210;  
Best Local Similarity 34.0%; Pred. No. 12;  
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;  
Oy 9 CSRGSSWADLKCMD-----CAS-----CRAPHSDFCLG-CAAAPAP 47  
Dy 194 CPNGSCWAGEENCQKTKIICAAQCSCRCRCGKSPDCDCHNQCAAGCTGP 243  
RESULT 8  
LMAL\_HUMAN  
ID LMAL\_HUMAN STANDARD; PRT; 3075 AA.  
AC P25391;  
DT 01-MAY-1992 (Rel. 22, Created)



DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LAMININ ALPHA-1 CHAIN PRECURSOR (LAMININ A CHAIN).  
GN LAMAL OR LAMA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91333420; PubMed=1714537;  
RA Haaparanta T., Uitto J., Ruoslahti E., Engvall E.;  
RT "Molecular cloning of the cDNA encoding human laminin A chain.";  
RL Matrix 11:151-160(1991).  
RN [2]  
RP SEQUENCE OF 1-2628 FROM N.A.  
RX MEDLINE=91364789; PubMed=2049067;  
RA Nissinen M., Vuolteenaho R., Boot-Handford R., Kallunki P.,  
RA Trygvsøen K.;  
RT "Primary structure of the human laminin A chain. Limited expression  
in human tissues";  
RL Biochem. J. 276:369-379(1991).  
RN [3]  
RP SEQUENCE OF 2397-3072 FROM N.A.  
RX MEDLINE=89280632; PubMed=2733383;  
RA Olsen D., Nagayoshi T., Fazio M., Peltonen J., Jaakkola S.,  
RA Sanborn D., Sasaki T., Kuivaniemi H., Chu M.L., Deutzmann R.,  
RA Timpl R., Uitto J.;  
RT "Human laminin: cloning and sequence analysis of cDNAs encoding A, B1  
and B2 chains, and expression of the corresponding genes in human  
skin and cultured cells.";  
RL Lab. Invest. 60:772-782(1989).  
CC -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ  
IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
CC -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND  
LAMININ-3 (S-LAMININ).  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR  
COMPONENT).  
CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
CC -!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.  
CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
CC -!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.  
CC -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; X58531; CAA41418.1;  
DR PIR; S14458; S14458.  
DR HSP; P02468; 1TLE.  
DR MIM; 150320;  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001886; LamNT.  
DR InterPro; IPR000034; Laminin\_B.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001791; Laminin\_G.  
DR Pfam; PF00052; laminin\_B\_2.  
DR Pfam; PF00053; laminin\_EGF; 15.  
DR Pfam; PF00054; laminin\_G; 5.

DR Pfam; PF00055; laminin\_Nterm; 1.  
DR PRINTS; P00011; EGF\_LAMININ.  
DR ProDom; P0002082; LamNT; 1.  
DR ProDom; P0003031; Laminin\_B; 2.  
DR SMART; SM00180; EGF\_Lam; 14.  
DR SMART; SM00001; EGF\_Like; 1.  
DR SMART; SM00281; LamB; 2.  
DR SMART; SM00282; LamG; 5.  
DR SMART; SM00136; LamNT; 1.  
DR PROSITE; PS00022; EGF\_1; 11.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 15.  
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 3075 LAMININ ALPHA-1 CHAIN.  
FT DOMAIN 18 269 LAMININ N-TERMINAL (DOMAIN VI).  
FT DOMAIN 270 516 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN  
V).  
FT DOMAIN 270 326 LAMININ EGF-LIKE 1.  
FT DOMAIN 327 396 LAMININ EGF-LIKE 2.  
FT DOMAIN 397 453 LAMININ EGF-LIKE 3.  
FT DOMAIN 454 502 LAMININ EGF-LIKE 4.  
FT DOMAIN 503 512 LAMININ EGF-LIKE 5 (N-TERMINAL).  
FT DOMAIN 517 708 LAMININ DOMAIN IV 1 (DOMAIN IV B).  
FT DOMAIN 709 1159 9 X LAMININ EGF-LIKE REPEATS (DOMAIN  
III B).  
FT DOMAIN 709 741 LAMININ EGF-LIKE 5 (C-TERMINAL).  
FT DOMAIN 742 790 LAMININ EGF-LIKE 6.  
FT DOMAIN 791 848 LAMININ EGF-LIKE 7.  
FT DOMAIN 849 901 LAMININ EGF-LIKE 8.  
FT DOMAIN 902 950 LAMININ EGF-LIKE 9.  
FT DOMAIN 951 997 LAMININ EGF-LIKE 10.  
FT DOMAIN 998 1043 LAMININ EGF-LIKE 11.  
FT DOMAIN 1044 1089 LAMININ EGF-LIKE 12.  
FT DOMAIN 1090 1149 LAMININ EGF-LIKE 13.  
FT DOMAIN 1150 1159 LAMININ EGF-LIKE 14 (N-TERMINAL).  
FT DOMAIN 1160 1361 LAMININ DOMAIN IV 2 (DOMAIN IV A).  
FT DOMAIN 1362 1555 3.5 X LAMININ EGF-LIKE REPEATS (DOMAIN  
III A).  
FT DOMAIN 1362 1402 LAMININ EGF-LIKE 14 (C-TERMINAL).  
FT DOMAIN 1403 1451 LAMININ EGF-LIKE 15.  
FT DOMAIN 1452 1508 LAMININ EGF-LIKE 16.  
FT DOMAIN 1509 1555 LAMININ EGF-LIKE 17.  
FT DOMAIN 1556 2125 DOMAIN II AND I.  
FT DOMAIN 2126 3075 5 X LAMININ G-LIKE REPEATS (DOMAIN G).  
FT DOMAIN 2140 2327 LAMININ G-LIKE 1.  
FT DOMAIN 2328 2509 LAMININ G-LIKE 2.  
FT DOMAIN 2510 2736 LAMININ G-LIKE 3.  
FT DOMAIN 2737 2913 LAMININ G-LIKE 4.  
FT DOMAIN 2914 3075 LAMININ G-LIKE 5.  
FT DOMAIN 1706 1796 COILED COIL (POTENTIAL).  
FT DOMAIN 1968 1989 COILED COIL (POTENTIAL).  
FT DOMAIN 2088 2120 COILED COIL (POTENTIAL).  
FT SITE 2534 2536 CELL ATTACHMENT SITE.  
FT DISULFID 270 279 BY SIMILARITY.  
FT DISULFID 272 290 BY SIMILARITY.  
FT DISULFID 292 301 BY SIMILARITY.  
FT DISULFID 297 305 POTENTIAL.  
FT DISULFID 304 324 BY SIMILARITY.  
FT DISULFID 327 336 BY SIMILARITY.  
FT DISULFID 329 361 BY SIMILARITY.  
FT DISULFID 364 373 BY SIMILARITY.  
FT DISULFID 376 394 BY SIMILARITY.  
FT DISULFID 397 409 BY SIMILARITY.  
FT DISULFID 399 427 BY SIMILARITY.  
FT DISULFID 429 438 BY SIMILARITY.  
FT DISULFID 441 451 BY SIMILARITY.  
FT DISULFID 454 467 BY SIMILARITY.  
FT DISULFID 456 471 BY SIMILARITY.  
FT DISULFID 473 482 BY SIMILARITY.  
FT DISULFID 485 500 BY SIMILARITY.  
FT DISULFID 742 751 BY SIMILARITY.









use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X91245; CAA62627.1; -  
EMBL; AF090832; AAC61591.1; -  
EMBL; AE003672; AAC54063.1; -  
HSSP; Q05158; 1QJ1.  
FlyBase; FBgn0014863; Mlp84B.  
InterPro; IPR001781; LIM.  
Pfam; PF00412; LIM; 5.  
ProDom; PD000094; LIM; 5.  
SMART; SM00132; LIM; 5.  
PROSITE; PS00478; LIM\_DOMAIN\_1; 5.  
PROSITE; PS50023; LIM\_DOMAIN\_2; 5.  
Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc; Myogenesis;  
Developmental protein; Differentiation.  
LIM 1.  
FT DOMAIN 12 63  
FT DOMAIN 65 80  
FT DOMAIN 66 71  
FT DOMAIN 120 172  
FT DOMAIN 175 180  
FT DOMAIN 178 189  
FT DOMAIN 222 274  
FT DOMAIN 276 291  
FT DOMAIN 325 377  
FT DOMAIN 379 390  
FT DOMAIN 421 473  
FT DOMAIN 475 490  
SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match 20.4%; Score 61.5; DB 1; Length 495;  
Best Local Similarity 31.1%; Pred. No. 7;  
Matches 14; Conservative 4; Mismatches 20; Indels 7; Gaps 1;

QY 10 SRGSSWSADLRCMDCAS-----CRARPHSDFCILGCAAAPAP 47  
DB 135 ARGSRWHKCFKCGTKCKGLDSTLCCEAPDKNIYKGCYAKKFGP 179

RESULT 12

MT3\_RAT  
ID MT3\_RAT STANDARD; PRT; 66 AA.  
AC P37361;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE METALLOTHIONEIN-III (MT-III) (GROWTH INHIBITORY FACTOR) (GIF).  
GN MT3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RX MEDLINE=94018480; PubMed=8412560;  
RA Kobayashi H., Uchida Y., Ihara Y., Nakajima K., Kohsaka S.,  
RA Miyatake T., Tsuji S.;  
RT "Molecular cloning of rat growth inhibitory factor cDNA and the  
expression in the central nervous system."  
RL Brain Res. Mol. Brain Res. 19:188-194(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Glial tumor;  
RA Anourex M.C., Rethaus E., Wurch T., Colpaert F.C., Pauwels P.J.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Chapman G.A., Kille P.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND  
ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).  
-!- TISSUE SPECIFICITY: BRAIN.  
-!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; S65838; AAB28366.1; -  
EMBL; X89603; CAA61762.1; -  
EMBL; Y08235; CAA69404.1; -  
HSSP; P18055; LMRB.  
InterPro; IPR003019; Metallthion.  
InterPro; IPR000006; Metallthion\_vert.  
Pfam; PF00131; metalthio; 1.  
PRINTS; PR00860; MTVERTEBRATE.  
PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;  
Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;  
Acetylation.

FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
FT DOMAIN 1 30 BETA.  
FT DOMAIN 31 66 ALPHA.  
FT METAL 6 6 CLUSTER B (BY SIMILARITY).  
FT METAL 8 8 CLUSTER B (BY SIMILARITY).  
FT METAL 14 14 CLUSTER B (BY SIMILARITY).  
FT METAL 16 16 CLUSTER B (BY SIMILARITY).  
FT METAL 20 20 CLUSTER B (BY SIMILARITY).  
FT METAL 22 22 CLUSTER B (BY SIMILARITY).  
FT METAL 25 25 CLUSTER B (BY SIMILARITY).  
FT METAL 27 27 CLUSTER B (BY SIMILARITY).  
FT METAL 30 30 CLUSTER B (BY SIMILARITY).  
FT METAL 34 34 CLUSTER A (BY SIMILARITY).  
FT METAL 35 35 CLUSTER A (BY SIMILARITY).  
FT METAL 37 37 CLUSTER A (BY SIMILARITY).  
FT METAL 38 38 CLUSTER A (BY SIMILARITY).  
FT METAL 42 42 CLUSTER A (BY SIMILARITY).  
FT METAL 45 45 CLUSTER A (BY SIMILARITY).  
FT METAL 49 49 CLUSTER A (BY SIMILARITY).  
FT METAL 51 51 CLUSTER A (BY SIMILARITY).  
FT METAL 62 62 CLUSTER A (BY SIMILARITY).  
FT METAL 64 64 CLUSTER A (BY SIMILARITY).  
FT METAL 65 65 CLUSTER A (BY SIMILARITY).  
SEQUENCE 66 AA; 6809 MW; BE7538E8564EBF8 CRC64;

Query Match 20.3%; Score 61; DB 1; Length 66;  
Best Local Similarity 35.0%; Pred. No. 1,4;  
Matches 14; Conservative 2; Mismatches 22; Indels 2; Gaps 2;

QY 4 PCTAFCSRGSSWSADLRCMDCASCARPHSDFCILGCAAA 43  
DB 3 PETCPCTGGCTCS-DKC-KCKGCKTCKCKSCCCCPA 40

RESULT 13

MT3\_MOUSE  
ID MT3\_MOUSE STANDARD; PRT; 68 AA.  
AC P28184;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE METALLOTHIONEIN-III (MT-III) (GROWTH INHIBITORY FACTOR) (GIF).  
GN MT3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:24 ; Search time 54.85 Seconds  
(without alignments)  
138.672 Million cell updates/sec

Title: US-09-742-454A-7\_COPY\_28\_79

Perfect score: 301

Sequence: 1 EQAPGAPCRSGSGWSADLD.....HSDFCLGCAAPAPFRLW 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 301   | 100.0       | 129    | 4 Q9NP84  | Q9np84 homo sapien |
| 2          | 276   | 91.7        | 129    | 11 Q9Q2W3 | Q9qzw3 mus musculu |
| 3          | 276   | 91.7        | 129    | 11 Q9CR75 | Q9cr75 mus musculu |
| 4          | 67.5  | 22.4        | 160    | 2 O86474  | O86474 rhodobacter |
| 5          | 67    | 22.3        | 94     | 4 Q9HCS0  | Q9hcs0 homo sapien |
| 6          | 66.5  | 22.1        | 420    | 5 P91776  | P91776 pacifastacu |
| 7          | 66    | 21.9        | 549    | 2 Q9HVT3  | Q9hvt3 pseudomonas |
| 8          | 64.5  | 21.4        | 112    | 11 Q9J148 | Q9ji48 mus musculu |
| 9          | 64.5  | 21.4        | 768    | 13 Q98TH8 | Q98th8 cyprinus ca |
| 10         | 64    | 21.3        | 115    | 4 Q9NZF1  | Q9nzf1 homo sapien |
| 11         | 64    | 21.3        | 473    | 11 Q9ESE0 | Q9ese0 rattus norv |
| 12         | 64    | 21.3        | 643    | 11 Q9ERV6 | Q9erv6 mus musculu |
| 13         | 64    | 21.3        | 655    | 11 Q9WVF5 | Q9wvf5 mus musculu |
| 14         | 64    | 21.3        | 669    | 4 O75441  | O75441 homo sapien |
| 15         | 64    | 21.3        | 1210   | 11 Q9EP98 | Q9ep98 mus musculu |
| 16         | 64    | 21.3        | 1511   | 4 O75412  | O75412 homo sapien |
| 17         | 64    | 21.3        | 1587   | 4 O00508  | O00508 homo sapien |
| 18         | 63.5  | 21.1        | 85     | 1 Q9P9E6  | Q9p9e6 methanococc |
| 19         | 63.5  | 21.1        | 708    | 13 P87363 | P87363 gallus gall |

#### ALIGNMENTS

RESULT 1

|        |   |              |      |         |
|--------|---|--------------|------|---------|
| Q9NP84 | Q9NP84  | PRELIMINARY: | PRT: | 129 AA. |
| AC     | Q9NP84;   |              |      |         |
| DT     | 01-OCT-2000 (Tremblrel. 15, Created)  |              |      |         |
| DT     | 01-OCT-2000 (Tremblrel. 15, Last sequence update)                           |              |      |         |
| DT     | 01-JUN-2001 (Tremblrel. 17, Last annotation update)                         |              |      |         |
| DE     | TYPE I TRANSMEMBRANE PROTEIN PRECURSOR (TYPE I TRANSMEMBRANE PROTEIN FN14). |              |      |         |
| DE     | FN14.   |              |      |         |
| GN     | FN14.   |              |      |         |
| OS     | Homo sapiens (Human).   |              |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;                     |              |      |         |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                  |              |      |         |
| OX     | NCBI_TaxID=9606;  |              |      |         |
| RN     | [1]   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |              |      |         |
| RA     | Tanaka S., Sugimachi K.;  |              |      |         |
| RT     | "Human homologue of Fn14.";   |              |      |         |
| RL     | Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.                     |              |      |         |
| RN     | [2]   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |              |      |         |
| RC     | TISSUE=PLACENTA;  |              |      |         |
| RX     | MEDLINE=20216634; PubMed=10751351;  |              |      |         |
| RA     | Feng S.-L.Y., Guo Y., Factor V.M., Thorgeirsson S.S., Bell D.W.,            |              |      |         |
| RT     | Testa J.R., Peiffley K.A., Winkles J.A.;                                    |              |      |         |
| RT     | "The Fn14 immediate-early response gene is induced during liver             |              |      |         |
| RT     | regeneration and highly expressed in both human and murine                  |              |      |         |
| RT     | hepatocellular carcinoma.";   |              |      |         |
| RL     | Am. J. Pathol. 156:1253-1261(2000).   |              |      |         |
| RN     | [3]   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |              |      |         |
| RC     | TISSUE=ENDOMETRIAL ADENOCARCINOMA;  |              |      |         |
| RA     | Strausberg R.;  |              |      |         |
| RL     | Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.                     |              |      |         |
| DR     | EMBL; AB035480; BAA94792.1; ..  |              |      |         |
| DR     | EMBL; AF191148; AAF69108.1; ..  |              |      |         |
| DR     | EMBL; BC002718; AAH02718.1; ..  |              |      |         |
| DR     | Signal; Transmembrane.  |              |      |         |
| KW     | POTENTIAL.  |              |      |         |
| FT     | SIGNAL 9 31   |              |      |         |
| SQ     | SEQUENCE 129 AA; 13911 MW; BF3DFB9C1E1C448 CRC64;                           |              |      |         |

Q92795 homo sapien  
Q9fca1 streptomyce  
Q9h2c9 homo sapien  
Q14226 homo sapien  
Q9gzx1 homo sapien  
P91404 caenorhabdi  
O27680 methanobact  
O24155 nicotiana t  
Q91jg5 mus musculu  
Q19088 caenorhabdi  
P92163 strongyloce  
O75136 homo sapien  
Q9qx70 rattus norv  
O75197 homo sapien  
Q9up66 homo sapien  
Q9ues7 homo sapien  
O75851 homo sapien  
Q9qzf2 mus musculu  
Q91310 pseudomonas  
O83022 rhodobacter  
Q9rad6 rhodobacter  
O95054 homo sapien  
Q9qee6 indian citr  
Q95407 homo sapien  
Q9uja2 homo sapien

```
Query Match      100.0%; Score 301; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 52
Db 28 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 79

RESULT 2
Q9Q2W3 PRELIMINARY; PRT; 129 AA.
ID Q9Q2W3;
AC Q9Q2W3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN FN14.
GN FGFRP2 OR FN14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RX MEDLINE=20020297; PubMed=10551889;
RA Meighan-Mantha R.L., Hsu D.K.W., Guo Y., Brown S.A.N., Feng S.-L.Y.,
RA Peifley K.A., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Richards C.W., Winkles J.A.;
RT "The mitogen-inducible Fn14 gene encodes a type I transmembrane
RT protein that modulates fibroblast adhesion and migration.";
RL J. Biol. Chem. 274:33166-33176(1999).
DR EMBL: AFI56164; AAF07882.1; -
DR MGD: MGI:1351484; Fgfrp2.
KW Transmembrane.
SQ SEQUENCE 129 AA; 13637 MW; 14B5C68BEF493385 CRC64;
```

```
Query Match      91.7%; Score 276; DB 11; Length 129;
Best Local Similarity 92.3%; Pred. No. 2.8e-25;
Matches 48; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 52
Db 28 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 79

RESULT 3
Q9CR75 PRELIMINARY; PRT; 129 AA.
ID Q9CR75;
AC Q9CR75;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FIBROBLAST GROWTH FACTOR REGULATED PROTEIN 2.
GN FGFRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA, AND CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
```

```
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005530; BAB24101.1; -
DR EMBL: AK005382; BAB23989.1; -
DR MGD: MGI:1351484; Fgfrp2.
SQ SEQUENCE 129 AA; 13641 MW; 1665C68B4D9A9253 CRC64;
```

Query Match 91.7%; Score 276; DB 11; Length 129;
Best Local Similarity 92.3%; Pred. No. 2.8e-25;
Matches 48; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 52
Db 28 EQPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLW 79

RESULT 4
O86474 PRELIMINARY; PRT; 160 AA.
ID O86474;
AC O86474;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NAF.
GN NAF.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI\_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 158;
RX MEDLINE=96332666; PubMed=8730872;
RA Reyes F., Roldan D., Klipp W., Castillo F., Moreno-Vivian C.;
RT "Isolation of periplasmic nitrate reductase genes from Rhodobacter
RT sphaeroides DSM 158: structural and functional differences among
RT prokaryotic nitrate reductases.";
RL Mol. Microbiol. 19:1307-1318(1996).
DR EMBL: Z46806; CAA86832.1; -
DR HSSP: P00195; 1CLF.
DR InterPro: IPR001450; 4FE4S\_ferredoxin.
DR Pfam: PF00037; fer4; 3.
DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; 3.
KW Iron-sulfur.
SQ SEQUENCE 160 AA; 16392 MW; 01FCADEF2EACB9EC6 CRC64;

Query Match 22.4%; Score 67.5; DB 2; Length 160;
Best Local Similarity 29.7%; Pred. No. 1;
Matches 19; Conservative 4; Mismatches 18; Indels 23; Gaps 2;

QY 3 APGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAA 43
Db 18 APTRPWR-----EADNARTGCAACACACGAGIVRMEAGLPQIAFAGTCFCGACAA 73

QY 44 PPAP 47
Db 74 CPAP 77

RESULT 5
Q9HCS0 PRELIMINARY; PRT; 94 AA.
ID Q9HCS0

|         |   |  |
|---------|---|--|
| DT      | 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)                   |  |
| DE      | HYPOTHETICAL PROTEIN PA4488.  |  |
| GN      | PA4488.   |  |
| OS      | Pseudomonas aeruginosa.   |  |
| OC      | Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;        |  |
| OC      | Pseudomonas.  |  |
| OX      | NCBI_TaxID=287;   |  |
| [1]     |   |  |
| RP      | SEQUENCE FROM N.A.  |  |
| RC      | STRAIN=PA01;  |  |
| RX      | MEDLINE=20437337; PubMed=10984043;                                    |  |
| RA      | Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,     |  |
| RA      | Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., |  |
| RA      | Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,   |  |
| RA      | Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,   |  |
| RA      | Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,          |  |
| RA      | Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;           |  |
| RT      | "Complete genome sequence of Pseudomonas aeruginosa PA01, an          |  |
| RT      | opportunistic pathogen.";   |  |
| RL      | Nature 406:959-964 (2000).  |  |
| RL      | EMBL; AE004862; AAG07876.1; -.  |  |
| KW      | Hypothetical protein; Complete proteome.                              |  |
| DQ      |   |  |
| SQ      | SEQUENCE 549 AA; 61710 MW; DAB72175DB036B7C CRC64;                    |  |
|         |   |  |
|         | Query Match 21.9%; Score 66; DB 2; Length 549;                        |  |
|         | Best Local Similarity 36.7%; Pred. No. 3.9;                           |  |
|         | Matches 18; Conservative 16; Mismatches 16; Indels 12; Gaps           |  |
| Qy      | 4 PGTPAGSGSSWADLKDCKMDCACRAPHSHDFCLGCAAPAPFRLIW 52                    |  |
|         | I : I I I : I             |  |
| Db      | 380 PASAPTRRIAWSADL----VLGATVNYHSD-----QPGPARLSW 416                  |  |
|         |   |  |
| RESULT  | 8   |  |
| Q9JJI48 | SEQUENCE FROM N.A.  |  |
| ID      | Q9JJI48   |  |
| AC      | PRELIMINARY; PRT; 112 AA.   |  |
| RC      | Q9JJI48;  |  |
| DT      | 01-OCT-2000 (TrEMBLrel. 15, Created)                                  |  |
| DT      | 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)                     |  |
| DT      | 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)                   |  |
| DE      | ONZIN (10 DAY OLD MALE PANCREAS CDNA, RIKEN FULL-LENGTH ENRICHED      |  |
| DE      | LIBRARY, CLONE:1810008C09, FULL INSERT SEQUENCE).                     |  |
| OS      | Mus musculus (Mouse).   |  |
| OC      | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |  |
| OC      | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.    |  |
| OX      | NCBI_TaxID=10090;   |  |
| [1]     |   |  |
| RP      | SEQUENCE FROM N.A.  |  |
| RC      | STRAIN=C57BL/6J; TISSUE=PANCREAS;                                     |  |
| RX      | MEDLINE=21085660; PubMed=11217851;                                    |  |
| RA      | Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,    |  |
| RA      | Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda Y.,    |  |
| RA      | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,    |  |
| RA      | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,    |  |
| RA      | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,       |  |
| RA      | Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,        |  |
| RA      | Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., |  |
| RA      | Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., |  |
| RA      | Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,      |  |
| RA      | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,      |  |
| RA      | Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,       |  |
| RA      | Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,   |  |
| RA      | Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,     |  |
| RA      | Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,          |  |
| RA      | Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., |  |
| RA      | Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., |  |
| RA      | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,     |  |

```
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AF263458; AAF76887.1; -
DR EMBL; AK007369; BAB24991.1; -
SQ SEQUENCE 112 AA; 12353 MW; 4FFBD5E8D91F0DA8 CRC64;

Query Match 21.4%; Score 64.5; DB 11; Length 112;
Best Local Similarity 34.1%; Pred. No. 1.7;
Matches 14; Conservative 2; Mismatches 24; Indels 1; Gaps 1;

Qy 4 PGTAPCSRGSSWSADLCKM-DCASCRARPHSDFCGLCAA 43
. || : || | | | | | | | | | | | | | | | |
Db 12 PGFVRAPQNSNWTSLCDSCFSGVCLGCTCFTCLGCOVA 52

RESULT 9
Q98TH8 PRELIMINARY; PRT; 768 AA.
AC Q98TH8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INTEGRIN BETA2-CHAIN.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura M., Fujiki K., Nakao M., Shin D., Yano T.;
RT "cDNA cloning of a carp homologue of mammalian integrin beta2-chain.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB013074; BAB39130.1; -
SQ SEQUENCE 768 AA; 84402 MW; 82564A9AA3ABDD65 CRC64;

Query Match 21.4%; Score 64.5; DB 13; Length 768;
Best Local Similarity 40.5%; Pred. No. 7.5;
Matches 17; Conservative 4; Mismatches 14; Indels 7; Gaps 3;

Qy 9 CSRGSSWSADLCKMDCASCRRAP--HSDFCGLGCAAAPAPF 48
. || | : || | | | | | | | | | | | | | | |
Db 600 CERGYSA----KCKTSSCQ-RPKESGSCVECLAFGTGPF 636

RESULT 10
Q9NZF1 PRELIMINARY; PRT; 115 AA.
AC Q9NZF1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BM-004.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhao M., Song H., Li N., Peng Y., Han Z., Chen Z.;
RT "A novel gene expressed in human bone marrow.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208846; AAF64260.1; -
SQ SEQUENCE 115 AA; 12507 MW; 24B6D7940390C6BF CRC64;

Query Match 21.3%; Score 64; DB 4; Length 115;
Best Local Similarity 30.8%; Pred. No. 2;
Matches 16; Conservative 3; Mismatches 23; Indels 10; Gaps 2;
```

```
Qy 2 QAP-----GTAPCSRGSSWSADLCKM-DCASCRARPHSDFCGLCAA 43
. || | : || | | | | | | | | | | | | | | |
Db 4 QAPVVVVTPQGVGPAPQNSNWTGMDCFSDCGVCLGCTGTCFPCPLGCOVA 55

RESULT 11
Q9ESE0 PRELIMINARY; PRT; 473 AA.
AC Q9ESE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR RELATED PROTEIN.
GN ERFP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GASTRO-DUODENAL MUCOUS;
RA Yu Y., Moshier J.A., Majumdar A.P.N.;
RT "Cloning of a novel EGFR-related peptide: A putative negative
RT regulator of EGFR.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187818; AAG17037.1; -
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR002174; Furin-like.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 1.
KW Receptor.
SQ SEQUENCE 473 AA; 52903 MW; 4E0D7C3074F28973 CRC64;

Query Match 21.3%; Score 64; DB 11; Length 473;
Best Local Similarity 31.7%; Pred. No. 5.9;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

Qy 4 PGTAP-----CSRGSSWSADLCKMD-----CAS-----CRARPHSDFCGLG-CAAAPAP 47
. || | : || | | | | | | | | | | | | | | |
Db 184. PSSCPKCDPSPGNGSGGGGEEQCQLTKIICAQQCRCRGRSPSDCHNQCAAGCTGP 243

RESULT 12
Q9ERV6 PRELIMINARY; PRT; 643 AA.
AC Q9ERV6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR ISOFORM 2.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, 129/SVEVTA;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Mahtie N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275366; AAG28046.1; -
DR EMBL; AF275364; AAG28046.1; JOINED.
DR EMBL; AF275365; AAG28046.1; JOINED.
DR MGD; MGI:95294; Egfr.
```

```

DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR SMART: SM00261; FU; 4.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 643 AA; 71477 MW; DEF22002C84911B1 CRC64;

Query Match 21.3%; Score 64; DB 11; Length 643;
Best Local Similarity 31.7%; Pred. No. 7.5;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

QY 4 PGTPAP-----CSRGSWSADLDKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47
DB 184 PSSCPKDPSPNGSCWGGGEENCQKLTKIIAQQSHRCGRSPSDCCNQCAGCTGP 243

RESULT 13
O9WVF5 PRELIMINARY; PRT; 655 AA.
AC O9WVF5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR (EPIDERMAL GROWTH FACTOR RECEPTOR
DE ISOFORM 3).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RA Reiter J.L., Threadgill D.W., Danielsen A.J., Schell C.,
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
RA Maible N.J.;
RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode
RT Carboxy-terminal Truncated Receptors.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV/TAC;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

```

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AF124513; AAD44149.1; -.
DR EMBL: AF275366; AAG28047.1; -.
DR EMBL: AF275364; AAG28047.1; JOINED.
DR EMBL: AF275365; AAG28047.1; JOINED.
DR EMBL: AK004944; BAB23688.1; -.
DR EMBL: AK004883; BAB23641.1; -.
DR EMBL: AK004911; BAB23662.1; -.
DR MGD: MGI:95294; Egrf.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR SMART: SM00261; FU; 3.
KW Receptor.
SQ SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;

Query Match 21.3%; Score 64; DB 11; Length 655;
Best Local Similarity 31.7%; Pred. No. 7.6;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

QY 4 PGTPAP-----CSRGSWSADLDKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47
DB 184 PSSCPKDPSPNGSCWGGGEENCQKLTKIIAQQSHRCGRSPSDCCNQCAGCTGP 243

RESULT 14
O75441 PRELIMINARY; PRT; 669 AA.
AC O75441;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LATENT TRANSFORMING GROWTH FACTOR-BETA BINDING PROTEIN 4 (FRAGMENT).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98325059; PubMed=9660815;
RA Saharinen J., Taipale J., Monni O., Keski-Oja J.;
RT "Identification and characterization of a new latent transforming
RT growth factor-beta-binding protein, LTBP-4.";
RL J. Biol. Chem. 273:18459-18469(1998).
DR EMBL: AF054502; AAC39882.1; -.
DR HSP: P35555; 1EMN.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR Pfam: PF00008; EGF; 6.
DR Pfam: PF00683; TB; 2.
DR SMART: SM00179; EGF_CA; 6.
DR SMART: SM00001; EGF_like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 6.
DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
DR PROSITE: PS01187; EGF_CA; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA; 71701 MW; 2C76C44BB11428A4 CRC64;

Query Match 21.3%; Score 64; DB 4; Length 669;
Best Local Similarity 35.6%; Pred. No. 7.7;
Matches 16; Conservative 3; Mismatches 14; Indels 12; Gaps 2;

```

